PRICE FIVE SHILLINGS.

The Institution of Electrical Engineers

ORIGINALLY

The Society of Telegraph Engineers

FOUNDED 1871

INCORPORATED BY ROYAL CHARTER 1921

SAVOY PLACE, VICTORIA EMBANKMENT, LONDON, W.C. 2.

Telegrams : " VOLTAMPERE, PHONE, LONDON." Telephone : TEMPLE BAR 7676.

PROCEEDINGS OF THE WIRELESS SECTION

OF THE INSTITUTION.

EDITED BY P. F. ROWELL, SECRETARY.

Vol. 13 JUNE, 1938. No. 38

(Issued two or three times yearly.)

CONTENTS.

(The Institution is not, as a body, responsible for the opinions expressed by individual authors or speakers.)

PAGE

Some Aspects of Magnetic Recording and its Application to Broadcasting.
A. E. BARRETT, M.A., and C. J. F. TWEED 73

A Theory of Fluctuation Noise .... D. A. BELL, B.A., B.Sc. 97

A Resonance Bridge for Use at Frequencies up to 10 Megacycles per Second.
Professor C. L. FORTESCUE, O.B.E., M.A., and G. MOLE, Ph.D., B.Sc. 112

Directional Recording of Radio Atmospherics .... F. E. LUTKIN 121

Discussion on " Theory and Performance of the Iconoscope " 135

Meetings 136

E. and F. N. SPON, LIMITED, 57 HAYMARKET, LONDON, S.W.1
CHEMICAL PUBLISHING COMPANY OF N.Y. INC., 148 LAFAYETTE STREET, NEW YORK.

The Institution of Electrical Engineers.

FOUNDED 1871

INCORPORATED BY ROYAL CHARTER 1921

PATRON: HIS MAJESTY KING GEORGE VI.

COUNCIL, 1937-1938.

President.

SIR GEORGE LEE, O.B.E., M.C.

Past-Presidents.

COLONEL R. E. CROMPTON, C.B., F.R.S.
SIR JAMES SWINBURNE, BART., F.R.S.
W. M. MORDEY.
SIR JOHN SNELL, G.B.E.
CHARLES P. SPARKS, C.B.E.
ROGER T. SMITH.
LL. B. ATKINSON.
J. S. HIGHFIELD.
F. GILL, O.B.E.
A. RUSSELL, M.A., D.Sc., LL.D., F.R.S.
W. B. WOODHOUSE.
W. H. ECCLES, D.Sc., F.R.S.
SIR ARCHIBALD PAGE.
LIEUT.-COL. K. EDGCUMBE, T.D.
COLONEL SIR THOMAS F. PURVES, O.B.E.
CLIFFORD C. PATERSON, O.B.E., D.Sc.
J. M. DONALDSON, M.C.
PROF. E. W. MARCHANT, D.Sc.
P. V. HUNTER, C.B.E.
PROF. W. M. THORNTON, O.B.E., D.Sc., D.Eng.
J. M. KENNEDY, O.B.E.
H. T. YOUNG.

Vice-Presidents.

SIR NOEL ASHBRIDGE, B.Sc. (Eng.).
J. R. BEARD, M.Sc.
A. P. M. FLEMING, C.B.E., D.Eng., M.Sc.
JOHNSTONE WRIGHT.

Honorary Treasurer.

W. MCCLELLAND, C.B., O.B.E.

Ordinary Members of Council.

T. E. ALLIBONE, D.Sc., Ph.D.
COL. A. S. ANGWIN, D.S.O., M.C.
E. S. BYNG.
T. CARTER.
C. E. FAIRBURN, M.A.
F. FORREST.
PROF. C. L. FORTESCUE, O.B.E., M.A.
P. GOOD.
E. M. LEE, B.Sc.
E. LEETE.

Ordinary Members of Council—(*Continued*).

S. W. MELSOM.
F. E. J. OCKENDEN.
P. L. RIVIÈRE.
P. J. ROBINSON, M.Eng.
C. RODGERS, O.B.E., B.Sc., B.Eng.
C. D. TAITE.
C. R. WESTLAKE.
G. A. WHIPPLE, M.A.

Chairmen and Past-Chairmen of Local Centres.

Argentine Centre:
C. G. BARKER.
*H. C. SIDDELEY.

China Centre:
S. FLEMONS.
*C. R. WEBB.

Irish Centre:
LT.-COL. K. E. EDGEWORTH, D.S.O., M.C.
*W. STOREY.

Mersey and North Wales (Liverpool) Centre:
G. K. PATON.
*F. E. SPENCER.

North-Eastern Centre:
W. D. HORSLEY.
*D. M. BUIST.

North Midland Centre:
J. G. CRAVEN.
*S. R. SIVIOUR.

North-Western Centre:
J. W. THOMAS, LL.B., B.Sc.Tech.
*G. L. PORTER.

Scottish Centre:
MAJOR H. BELL, O.B.E., T.D.
*H. C. BABB.

South Midland Centre:
H. HOOPER.
*E. A. REYNOLDS, M.A.

Western Centre:
H. G. WEAVER.
*J. W. SPARK.

* Past-Chairman.

Chairmen of Sections.

H. COBDEN TURNER (*Meter and Instrument*).
J. L. EVE (*Transmission*).
T. WADSWORTH, M.Sc. (*Wireless*).

Secretary.

P. F. ROWELL.

Assistant Secretaries.

F. W. HEWITT, M.Sc.; A. E. RAYNER.

WIRELESS SECTION COMMITTEE, 1937-1938

Chairman: T. WADSWORTH, M.Sc.
Vice-Chairman: W. J. PICKEN.
Immediate Past-Chairman: E. MALLETT, D.Sc.(Eng.).

J. R. BEARD, M.Sc. (nominated by the Council).
H. BISHOP, B.Sc.(Eng.).
W. J. BROWN, B.Sc.
S. BRYDON, D.Sc.
R. P. G. DENMAN, M.A.
W. T. DITCHAM.
COL. R. ELSDALE, O.B.E., M.C. (nominated by the War Office).
A. J. GILL, B.Sc.(Eng.) (nominated by the Post Office).
CAPT. P. F. GLOVER, R.N. (nominated by the Admiralty).
T. E. GOLDUP.
N. F. S. HECHT (nominated by the Air Ministry).
H. L. KIRKE.
SIR GEORGE LEE, O.B.E., M.C. (*President, I.E.E.*).
E. B. MOULLIN, M.A.
A. H. MUMFORD, B.Sc.(Eng.).
S. B. SMITH.
R. L. SMITH-ROSE, Ph.D., D.Sc
W. URE, B.Sc.

And the Chairman of the I.E.E. Papers Committee.

SOME ASPECTS OF MAGNETIC RECORDING AND ITS APPLICATION TO BROADCASTING*

By A. E. BARRETT, M.A., and C. J. F. TWEED, Associate Members.

(*Paper first received* 17*th December*, 1936, *and in final form* 6*th September*, 1937; *read before the* Wireless Section 1*st December*, 1937.)

SUMMARY

This paper presents a somewhat simplified consideration of the processes occurring in the magnetic recording of sound on steel tape, and describes some of the apparatus and methods used in this country in its application to broadcasting.

A brief historical account is given of the early uses of magnetic recording. The essential processes involved are then discussed in a simplified form, and the effects of the finite longitudinal spread of the recording field, and of the self-demagnetization of the tape, are considered before the question of background noise is introduced.

In the next section certain fundamental experiments are described, followed by a brief account of various possible methods of saturating, recording, and reproducing. Considerations relating to the conditions for the optimum working of any particular arrangement are put forward, with the results of tests of two of those which are now in use. Some details are given concerning the tape itself, after which a recording machine is described.

The paper concludes with a short account of the technical and programme service requirements for a recording system for use in connection with broadcasting, and a description of a recording channel at the British Broadcasting Corporation's premises at Maida Vale.

SECTION (1)

Introduction

Very little information has been published in this country on the subject of magnetic recording on steel tape. The object of this paper is to give an outline of the theory, and to describe some of the methods and apparatus that have been developed by the British Broadcasting Corporation and Marconi's Wireless Telegraph Co. during recent years.

A certain amount of the elementary theory has been given in previous papers published in other countries, and some of this ground has been covered again. Apart from this, the scope of the paper is limited to work that has been carried out in this country alone.

The research and development work on the subject is by no means complete, but a standard of performance is now obtainable which enables this system of recording to be extensively used in connection with broadcasting.

History

The process of magnetic recording was introduced by Poulsen under the name of the Telegraphone, and was used in conjunction with his arc system of radio-telegraphy. Telegraph signals were transmitted at high hand speed and were recorded at the receiving station on a steel wire which was driven at a higher velocity when recording than during the subsequent reproducing process. Thus the signals could be transcribed by the receiving operator at a convenient speed. In this application, variations in the velocity of the steel wire, and the effects of frequency or amplitude distortion, were relatively unimportant. The recording and reproducing electromagnetic systems were somewhat crude when compared with present-day apparatus, but even at the early date of 1900, when the system was developed, Poulsen clearly proved the possibilities of the magnetic recording of sound, although suitable amplifiers employing thermionic values were not available.

About the year 1924 a German engineer, Dr. Stille, commenced his research on the subject. His work was entirely devoted to the problems associated with the recording of sound, and he made a careful study of both the electromagnetic and the mechanical problems involved. The greatest advantages accruing from Dr. Stille's work were, first, the introduction of steel tape instead of wire, which overcame certain magnetic and mechanical difficulties, and, secondly, the development of an electromagnetic system which resulted in a reduction of wave-form distortion.

The development of magnetic recording in this country dates from the year 1930, when the British Broadcasting Corporation first tested the system under the name of the Blattnerphone. It was soon recognized that, in order to attain a standard of recording suitable for broadcasting purposes, the general performance of the equipment would have to be considerably improved. Experimental work was carried out in collaboration with British Blattnerphone, Ltd., and this resulted in a marked improvement in the quality of reproduction. When the Empire broadcasting service was inaugurated, the system proved to be very useful for recording extracts from the home programmes for subsequent reproduction in the Empire transmissions. The first important programme to be radiated to the Empire by this means was that of Christmas Day, 1932, which included the speech of His late Majesty King George V.

Early in 1933 Marconi's Wireless Telegraph Co. designed a machine in which many of the mechanical difficulties experienced with earlier models have been overcome. Development work has been carried out by both the B.B.C. and the Marconi Co.; this has resulted in a further considerable improvement in the performance of the system.

Fundamental Processes

The process of magnetic recording depends fundamentally upon two properties of steel; remanence and coercivity. It is by virtue of the property of coercivity that neighbouring elements in a length of steel can be magnetized, to a great extent, independently.

---

* Reprinted from *Journal I.E.E.*, 1938, vol. 82, p. 265.

Magnetic recording in its simplest form can be carried out by passing a length of steel tape or wire at constant velocity through an electromagnetic device (a recording head) which will produce variations in the distribution of remanent flux along its length.

Reproduction can be effected by passing the magnetized tape at an identical velocity through a device (a reproducing head) which is sensitive to changes of the flux in the tape. The spatial variations of flux will then be transformed into alternations of e.m.f. with respect to time. In the ideal case these correspond to the original alternations applied to the recording head.

It is necessary that any previous variations in the flux distribution in the tape should be removed before recording. This can be done by passing the tape through a " wiping head," which magnetizes the tape to saturation.

The essentials of the system are thus:—

(1) A length of steel tape driven at constant velocity.
(2) A wiping head.
(3) A recording head.
(4) A reproducing head.

In the simplest case the heads might take the form of coils wound round the tape, but if consideration is given to the wide field distribution associated with such coils it is clear that it would be impossible to obtain satisfactory results with practicable tape velocities, except at very low frequencies. It is therefore necessary to introduce a medium of high permeability into the system in order to concentrate the field.

The heads designed to perform these various functions are similar in appearance and construction; they are shown in Figs. 19 and 20 (Plates 1 and 2). Each head consisists of two similar blocks of insulating material held together by a hinge at the back. When closed a slot is formed for the tape to pass through. In each half a strip of magnetic alloy known as a " pole piece " may be placed, making mechanical contact with the tape when the head is closed. The pole-pieces are at right angles to the plane of the tape and are surrounded by coils through which they are free to slide, the pressure of the pole-pieces on the tape being adjusted by spring-loaded plungers. One half of the head can be moved in relation to the other, in the direction of the movement of the tape, so that a variable longitudinal gap is formed between the two pole-pieces.

The wiping, recording, and reproducing heads used in the experiments described have different types of coils and pole-pieces, but are in other respects identical.

The tape may be magnetized longitudinally or transversely, through its width or through its thickness, or in any combination of these directions. It is generally accepted that transverse magnetization of the tape through its width gives inferior results. This paper is concerned essentially with longitudinal magnetization, though it is inevitable that there will also be a transverse component through the thickness of the tape.

It is convenient in the first place to consider the various processes in a simplified form. Let it be assumed that the whole width of the tape is used, and that at any point in its length the flux distribution is substantially uniform over the cross-section.

The wiping head carries a direct current adequate to saturate the tape completely, so that it leaves the head fully magnetized, i.e. with the maximum remanent induction ($B_{r\,max.}$) in a longitudinal direction.

The Recording Process.

The object of the recording head is to produce variations in the flux in the tape of the largest amplitude that is consistent with linearity. It will be shown that the alternating current representing the programme to be recorded must be superimposed on a direct current, the purpose of which is to bring the tape to a suitable magnetic state. It is also necessary for the net current in the head to produce a field that is always opposed in direction to that of the wiping head or partial rectification may occur. For this reason the direct current in the recording head is usually referred to as the demagnetizing current.

Each elementary length of tape passing through the head will be demagnetized to an extent which is largely dependent upon the net value of the alternating and the steady components of the field acting upon it, and will leave the head with a corresponding value of remanent induction (B_r). If the recording current is sinusoidal, of frequency f cycles per sec., and the tape speed is u, variations in the flux in the tape will be produced of wavelength given by

$$\lambda = \frac{u}{f}$$

The process may be followed by referring to a ($B - H$) curve for the tape (Fig. 1). An elementary length of tape enters the wiping head in some unknown condition, and is then brought to saturation (point 1). It leaves the wiping head fully magnetized (point 2) and is then subjected to the recording field. If, at this moment, the sinusoidal a.c. component is zero, the effect of the d.c. component is to reduce the magnetization to a state corresponding to point 3, and on leaving the head the remanent magnetization is represented by point 3a. The effect of the a.c. component is to increase and decrease the effective value of the demagnetizing current. If the a.c. component is at either of its peak values the tape is demagnetized to a degree represented by points 4 or 5, and leaves the head in a state corresponding to points 4a or 5a. Point 3 may be called the working point, and its position is determined by the state of the tape as it enters the recording head, and by the value of the d.c. component of the demagnetizing current.

It is important to realize that the whole of the tape is not subjected to an alternating cycle, but that each elementary length of tape is demagnetized separately to a degree depending upon the net demagnetizing current, so that the tape leaves the head with a value of B_r (the remanent induction) which varies sinusoidally along its length, provided that there is a linear relation between the demagnetizing current and B_r. Referring again to Fig. 1, this implies that, in this simplified consideration, variations in induction represented by points 4a, 3a,

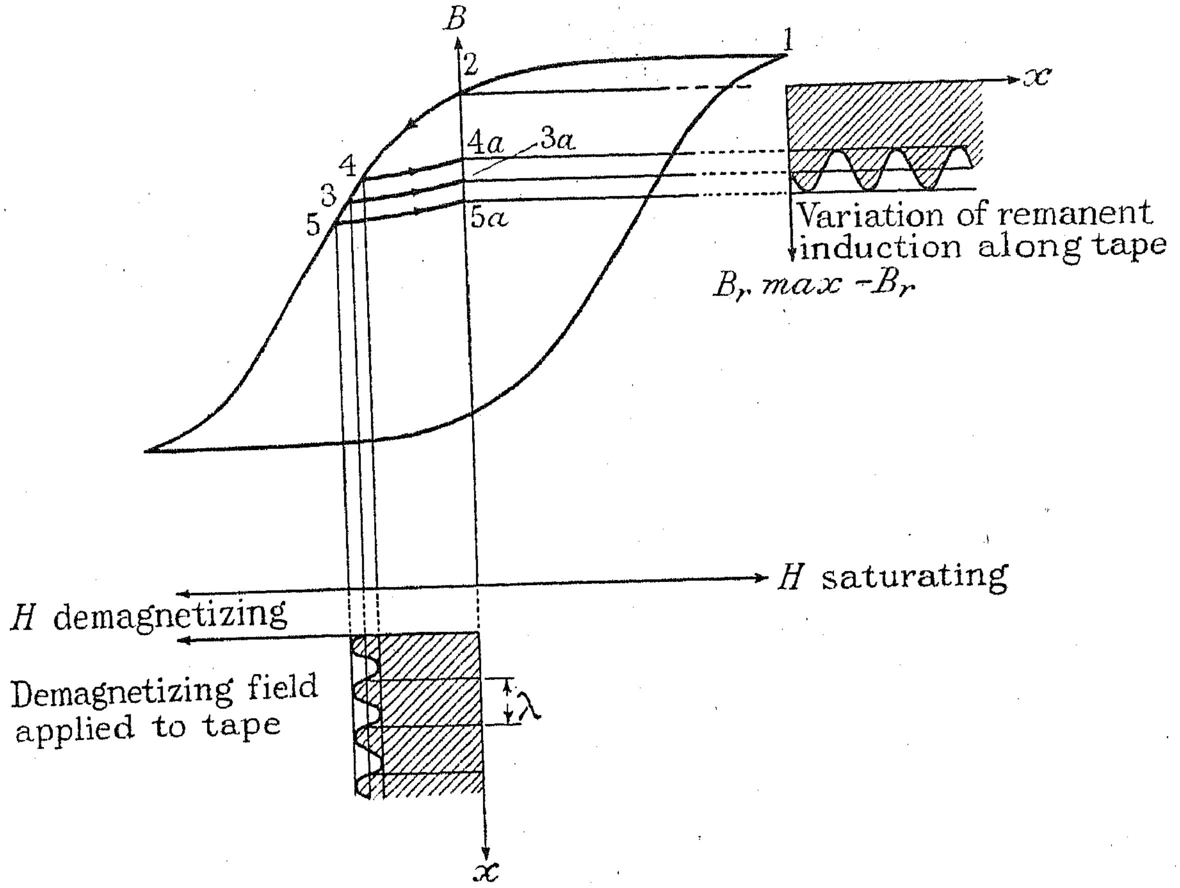
B
1
2
x
4a
3a
4
3
5
5a
Variation of remanent induction along tape
B_r max $-B_r$
H saturating
H demagnetizing
Demagnetizing field applied to tape
λ
x

Fig. 1.—Simplified recording process.

and 5a (rather than 4, 3, and 5), should be proportional to the corresponding values of the demagnetizing current. Now in order that points 4, 3, and 5, should have a linear relationship to the demagnetizing current through as wide a range as possible, the largest possible hysteresis loop must be used. In other words, the tape should enter the recording head fully magnetized. In practice, it appears that this condition also fulfils the requirement that the range through which the remanent magnetization is linear with the demagnetizing current should be a maximum. The optimum value of the d.c. component of the demagnetizing current to give the maximum variation of remanent flux with minimum distortion can be found empirically.

Let it be assumed that the variation of longitudinal flux in the tape after it has left the recording head is sinusoidally distributed along its length. Any variation in the flux in the tape implies that a corresponding number of lines of magnetic induction enter or leave the tape. These lines may be termed the external flux. It follows that the surface density (B_s) of external flux at any point in the length of the tape is proportional to the rate of change of the flux along the tape at that point. This surface flux density is independent of the mean state of magnetization of the tape, which is the datum about which variations take place, and which is determined by the d.c. component of the recording field. There is one value of this component for which variations will occur about the state of zero longitudinal magnetization as a datum, but in practice a somewhat lower value is used, so that the direction of magnetization in the mean state is the same as that of the fully magnetized tape.

In this consideration it may be assumed that the value of this datum has no effect upon the external flux distribution, and it is taken as being zero. A diagrammatic representation of the longitudinal and external flux is shown in Fig. 2(*b*). The value of the longitudinal induction at any point is given by the width of the shaded area at that point, whilst its direction is given by the white arrow in the centre. If ϕ represents the longitudinal flux in the tape in any section and Φ is the

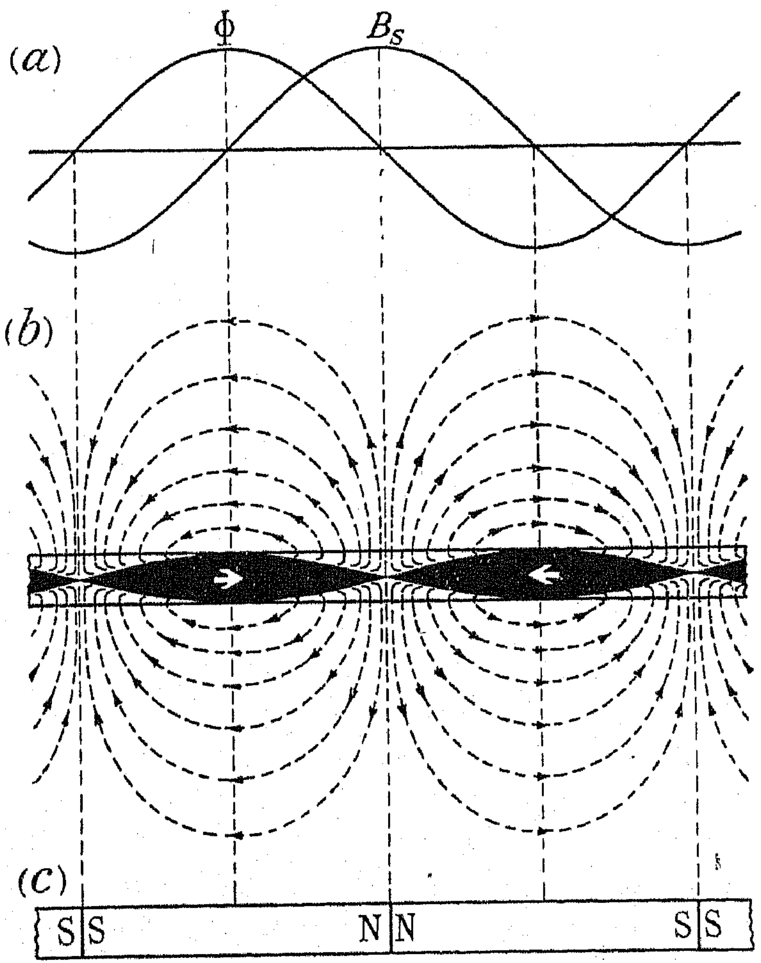

Fig. 2.—Flux distribution for recorded tape.

peak value of this, then $\phi = \Phi \sin 2\pi x/\lambda$, where x represents distance along the tape and the mean surface flux density B_s is proportional to $d\phi/dx$, which is given by

$$\frac{d\phi}{dx} = \frac{2\pi}{\lambda}\Phi \cos\frac{2\pi x}{\lambda}$$

The distribution of external flux is thus in phase quadrature with the longitudinal flux (Fig. 2*a*), and its maximum values are reached where the rate of change of longitudinal flux is highest, though the flux itself is at its datum value. At these points poles will, in effect, be formed on the surface of the tape. From the above expression it will be seen that, for a given value of Φ, the strength of these surface poles is inversely proportional to the wavelength. Since the tape moves with constant velocity $u = x/t$ the rate of change of the flux along the tape can be expressed as a time rate of change of flux passing a given point. Thus

$$B_s \propto \frac{\omega}{u}\Phi \cos \omega t, \text{ where } \omega = 2\pi f.$$

In an ideal case, the external flux passing any given point will therefore vary sinusoidally with time, and the problem now arises of transforming these variations of flux back into variations of e.m.f.

The Reproducing Process.

An examination of the reproducing process is necessarily very complicated, since it should take into account the effect of the pole-pieces upon the flux distribution in and around the tape, for in any system in which pole-pieces are used the external flux distribution of the tape is bound to be somewhat altered by their presence. Fortunately, however, it is not inevitable that the reproduced waveform will be correspondingly distorted. It has already been explained that some form of iron circuit is essential, since the physical dimensions of any coil used for reproducing purposes will be large compared with the wavelength at the higher frequencies. A relatively large coil is, of course, necessary in order that the reproduced e.m.f. should be high compared with the valve noise produced by the first valve of the reproducing amplifier.

The surface poles on the tape due to the recorded signal will induce poles of opposite polarity at the tip of any pole-piece in contact with the tape, and like poles distributed along the surfaces and the far end of the pole-piece. If a coil is wound closely round the pole-piece, a certain percentage of the turns of this will be threaded by the flux that enters the pole-piece, and an e.m.f. proportional to the time-rate-of-change of flux-turns will be generated in the coil.

In the simplest case, for long wavelengths, the flux threading the coil would be proportional to the value of B_s, so that the reproduced voltage would be given by

$$v \propto \frac{dB_s}{dt} \propto \frac{\omega^2}{u}\Phi \sin \omega t$$

As the frequency is increased, however, the percentage of turns threaded is reduced, and, further, the effective magnetic permeability of the pole-piece metal becomes lower. Thus the efficiency of the reproducing system tends to fall off very rapidly as the frequency rises. Furthermore, it is not possible for Φ to be maintained constant throughout the frequency range of the system, owing to the limitations of the recording head and to the effects of self-demagnetization. Thus an overall frequency characteristic that rises steeply as the frequency increases is not obtained, though it might be expected from the above expression for the reproduced voltage.

The simple explanation which has just been put forward serves to give some general idea of the principles of magnetic recording. In practice the recording process is very much more complicated, and, while it would be impossible to give even a qualitative analysis in detail, there are two important factors which deserve consideration.

The first is the effect of the finite longitudinal spread of the recording field, and the second is the self-demagnetizing effect of the surface poles formed by the recorded signal.

Effect of Recording with a Finite Field Distribution.

So far in this consideration the effective longitudinal spread of the recording field has not been taken into account. This must be finite in order to produce variations in the longitudinal flux in the tape. On the other hand, if the effective spread is comparable with the wavelength of the flux variations at any frequency, distortion of the waveform will result.

In order to simplify the explanation of this, an ideal case may be considered, in which it is assumed:—

(1) That the time of magnetization is effectively zero. This implies that the remanent magnetization of the tape after the magnetizing field has been removed is independent of the time during which this field was applied.

(2) That the final state of remanent magnetization in an elementary length of tape that has been saturated and then partially demagnetized depends only upon the peak value of the applied demagnetizing field, so that the subsequent application of a demagnetizing field of a lower value has no effect.

(3) That the value of this remanent magnetization bears a linear relationship to the current in the windings of the recording head, and that each elementary length of tape can be magnetized independently.

(4) That self-demagnetization effects can be neglected.

In the first case, let the spatial distribution of the applied demagnetizing field be rectangular, so that it reaches a maximum value instantaneously, is constant throughout a distance dx measured along the tape, and falls away instantaneously to zero. Let this value vary sinusoidally with time, about an arbitrary datum level.

Referring to Fig. 3, the tape is first of all saturated to state 1, and enters the influence of the recording field in state 2. There will always be a section of tape of length dx which is under the influence of the demagnetizing field. The distortion of the recorded waveform that results from the finite value of dx can best be visualized if it is imagined that an attempt is made to draw a sinusoidal curve with a very broad pen-nib. Although any one point on the nib will trace out a true sine wave, the outer edges of the curve so drawn will not be sinusoidal.

Two curves (A and B in Fig. 3) have been drawn which correspond to the demagnetizing effect of the extreme edges of the field distribution. The boundary curve of the shaded area will then give the peak values of the demagnetizing field which is applied to the tape at any point. The resulting changes in remanent induction ($B_{r\,max.} - B_r$) are shown by curve C, which thus gives the distribution of remanent induction along the tape. Clearly, unless dx is very small compared with the wavelength, so that the boundary curve is approximately sinusoidal, distortion due to even harmonics will occur. This form of distortion is dependent upon frequency and cannot be eliminated by reduction of amplitude. In order to minimize it, the length (dx) over which the field is maintained constant should be small compared with the shortest wavelength.

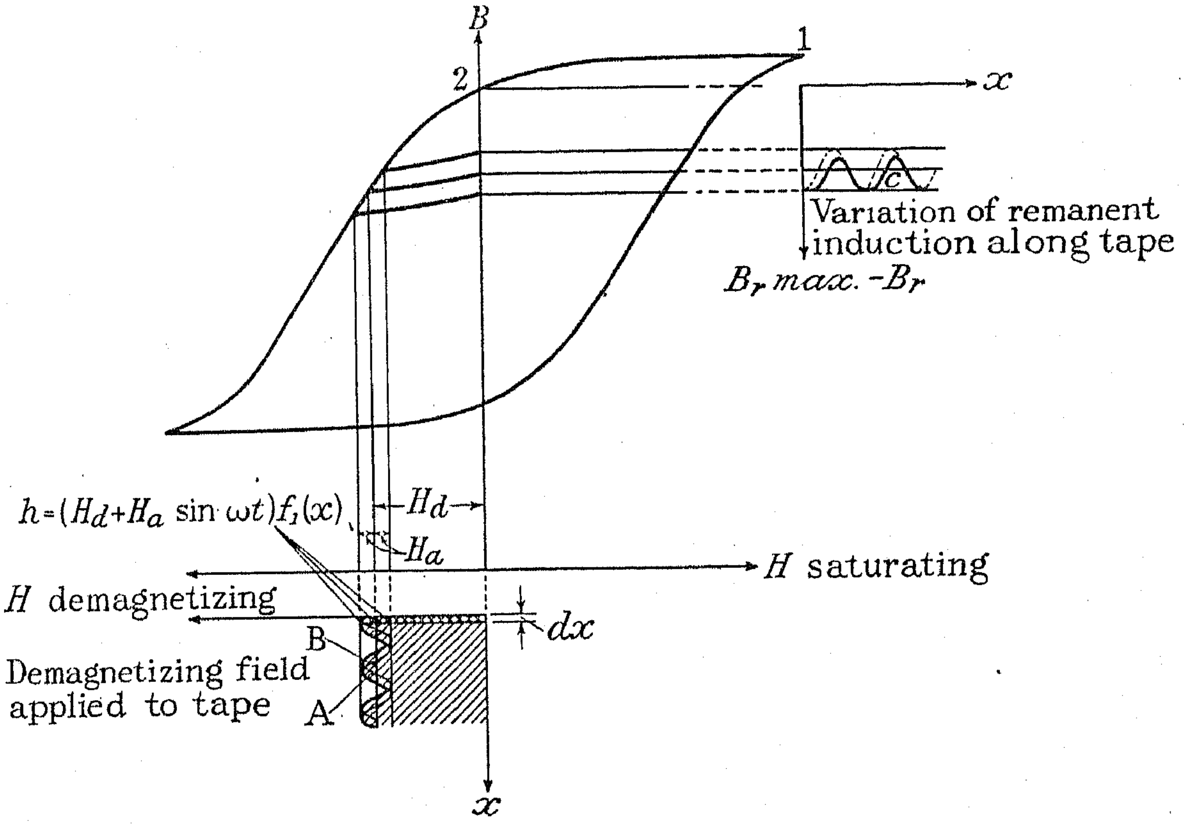
B
1
2
x
Variation of remanent induction along tape
$B_r\,max. - B_r$
$h = (H_d + H_a \sin \omega t) f_1(x)$
H_d
H_a
H saturating
H demagnetizing
dx
B
Demagnetizing field applied to tape
A
c
x

Fig. 3.—Distortion due to recording field distribution. $h = (H_d + H_a \sin \omega t)\, f_1\,(x)$.

In practice it is not possible to obtain a field distribution that falls away instantaneously to zero along the tape, and a second case will now be considered in which the demagnetizing field rises sharply to its maximum value and falls away upon some arbitrary curve. The value of the field represented by each point on the curve will vary sinusoidally in time about a datum fixed by the value of the d.c. component of the demagnetizing current.

If the field distribution is such that the reduction in field along the tape is not sufficiently rapid, an elementary length of tape which was originally demagnetized by a low value of field may subsequently be subjected to a higher value. Referring to Fig. 4, it will be seen that the boundary curve of the shaded portion indicates the actual peak value of demagnetizing field to which the tape will be subjected at any point in its length, and this will produce a distribution of remanent induction corresponding to curve D. This type of distortion is dependent upon both frequency and amplitude, and can be eliminated by a sufficient reduction of amplitude or increase in the wavelength.

In practice, the distortion due to both these types of field distribution may occur, but since it will be most pronounced at the higher frequencies the most serious harmonics are likely to be outside the useful range of the system. Intermodulation tones due to amplitude non-linearity, however, may still be reproduced.

Effect of Self-Demagnetization.

A length of tape, in which sinusoidal variations of flux have been recorded, can be represented by a series of magnets each half a wavelength long, placed end to end with like poles adjacent, as shown in Fig. 2(c). Since the self-demagnetization effect for any magnet increases as the length is reduced, for a constant cross-section, the effective pole strength is reduced as the recorded frequency is increased. With the normal tape speed of 1·5 metres per sec., the wavelength corresponding to a signal frequency of 5 000 cycles per sec. is 0·3 mm., so that each of these magnets will be 0·15 mm. long. As the thickness of the tape is 0·08 mm. and its width 3 mm., the self-demagnetization effect may well be of fundamental importance, especially at the higher frequencies. When this is taken into account, a new condition for amplitude linearity must be introduced. Referring to Fig. 5, the elements of tape which have been demagnetized by the applied demagnetizing field to states corresponding to points 3, 4, and 5, on the hysteresis loop would return to states corresponding to points 3a, 4a, and 5a, if the field acting upon them were reduced to zero. But the statement that there is an alternating external flux distribution about the tape and that, in effect, poles are formed on its surface, implies that the

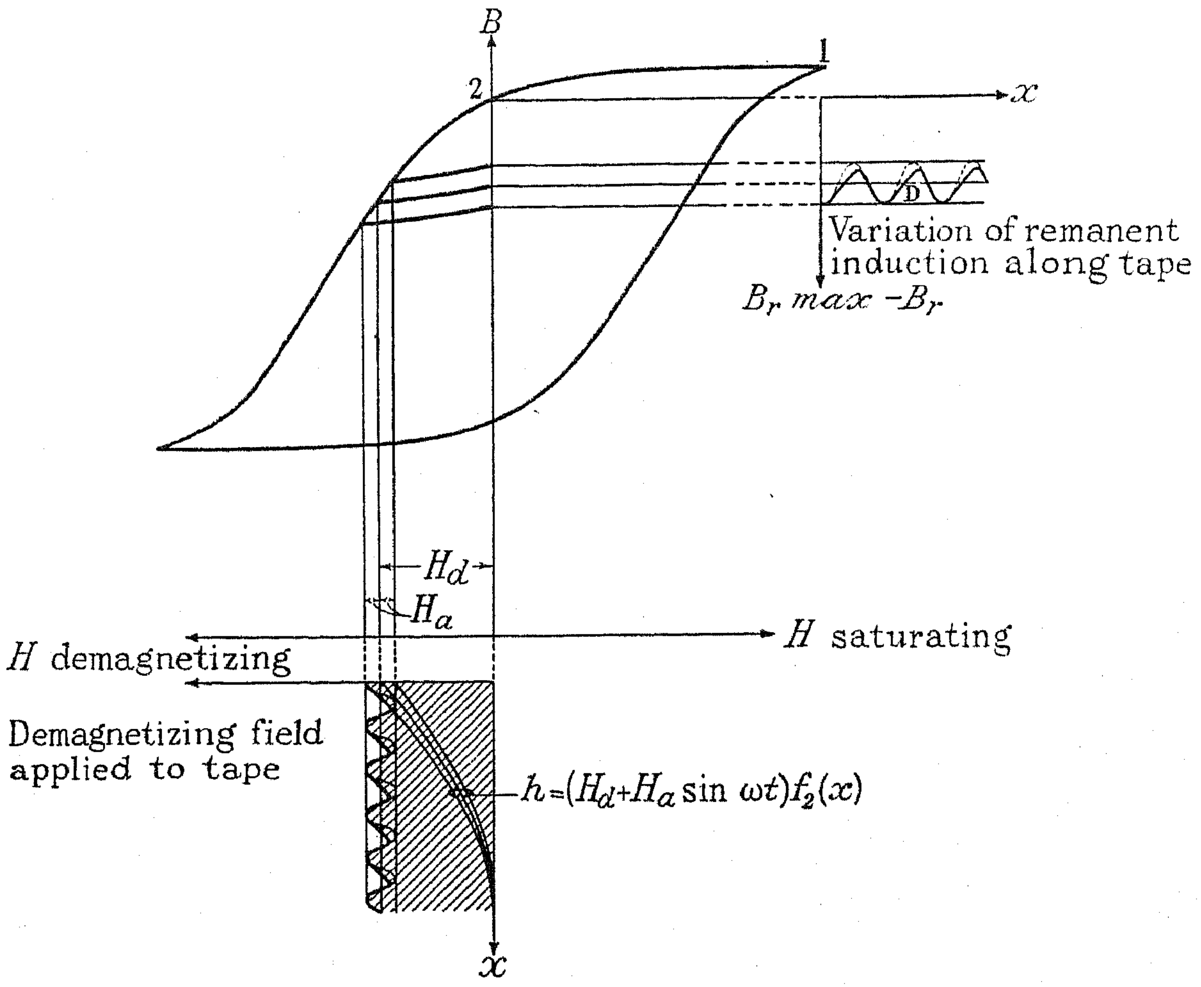
B
1
2
x
D
Variation of remanent induction along tape
B_r max $-B_r$
H_d
H_a
H saturating
H demagnetizing
Demagnetizing field applied to tape
$h=(H_d+H_a \sin \omega t)f_2(x)$
x

Fig. 4.—Distortion due to recording field distribution. $h = (H_d + H_a \sin \omega t) f_2(x)$.

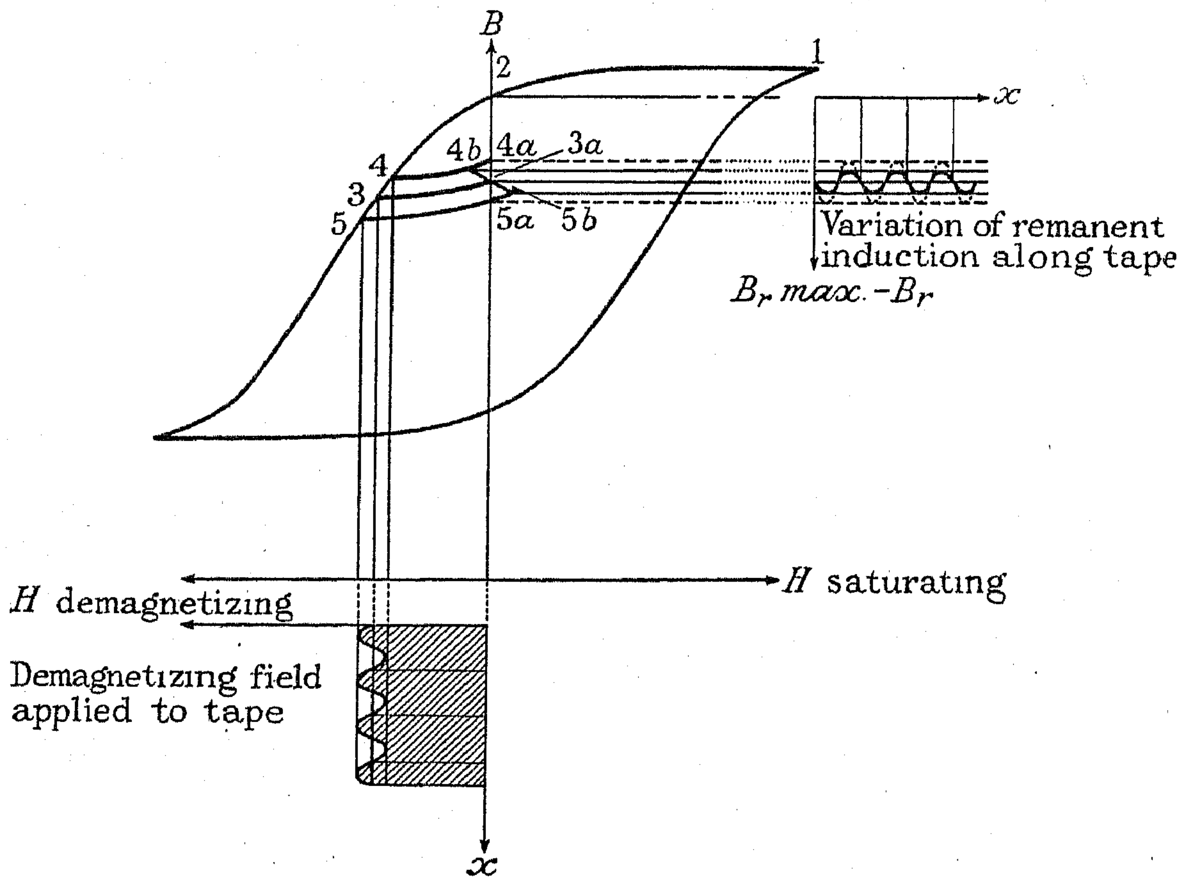
B
2
1
x
4
3
5
4b
4a
3a
5a
5b
Variation of remanent induction along tape
B_r max. $-B_r$
H saturating
H demagnetizing
Demagnetizing field applied to tape
x

Fig. 5.—Self-demagnetizing effect in recorded tape.

tape is subjected to a self-demagnetizing field that alternates along its length. It is impossible to differentiate between the effect of this self-demagnetizing field and that of an externally applied field upon an elementary cube inside the tape, so that it becomes necessary to modify the conception of the process represented on a $(B - H)$ characteristic.

The self-demagnetizing force acts in such a direction as to reduce the flux variations about the datum, increasing the demagnetization of the element which was subjected to the minimum recording field, and vice versa. The final points reached will then be 3a, 4b, and 5b, in Fig. 5, rather than 3a, 4a, and 5a. It is possible that for very long wavelengths, where the self-demagnetization is comparatively low, these sets of points may be almost coincident, but for the shorter wavelengths this will not be the case. Assuming a sinusoidal recording current, the new condition to be fulfilled for amplitude linearity in recording is that on the $(B - H)$ curve, point 3a should lie on a straight line midway between points 4b and 5b. From this point of view it would appear that it is even more important that the tape should have high coercivity than that the maximum value of remanent magnetism should be high.

This self-demagnetization effect, which is essentially the effect of the poles formed on the surface of the tape, still further complicates the process of recording. For, at the moment of recording, the tape is subjected to the resultant of the field due to these poles formed by the signal already recorded, and the applied field due to the recording head. For any given frequency the former will be alternating in value and in general will be opposed to, but not in exact antiphase with, the latter. At the higher frequencies these fields may be of the same order of magnitude, in which case the tape is subjected to a considerably lower net recording field than that estimated from the constants of the recording head.

Background Noise.

It can be seen that there are a number of conditions that must be fulfilled in order to avoid distortion in the recording process, and these determine the maximum amplitudes that can be recorded.

A lower limit to the permissible recording level is imposed by background noise, which is inherent in the use of any sound-recording medium yet developed. The reproduced level of background noise is mainly dependent upon the following factors:—

(*a*) Variations in the magnetic and mechanical properties of the tape itself, due to such considerations as granular structure and size, chemical composition, physical dimensions, and surface polish. Since the material of which the tape is composed is a mixture as distinct from an alloy, the above properties will largely depend upon the degree of uniformity which it is possible to maintain during the process of manufacture. Variations in the properties of the tape may also be produced by uneven stresses occurring during the operation of the machine.

(*b*) The mean state of magnetization of the tape leaving the recording head.

(*c*) Relative motion between the pole-pieces and the [illegible] and between the coils and the pole-pieces. This is largely due to unevenness of the tape surface and to mechanical vibration. It should not cause noise in the case of the wiping pole-pieces unless the final saturating flux is inadequate to saturate the tape completely. In the case of the recording and reproducing pole-pieces, the noise produced will depend upon the magnitude of the motion and upon the resulting change in the reluctance of the magnetic circuit involved.

(*d*) Variations in the magnetic properties of the pole-piece due to vibration when in contact with the moving tape.

(*e*) The relation between frequency and the recorded flux variations. This varies with the recording frequency characteristic used, and affects the noise level because it determines the frequency response of the reproducing amplifier.

(*f*) It is possible that the phenomenon of magnetostriction and the Barkhausen effect may contribute to the noise reproduced, but it is thought that their influence is probably small compared with the other factors enumerated.

It can be seen that the reproduced noise level is very much dependent upon the manufacture of the tape itself, and, while considerable improvement has been made during recent years, it is hoped that yet further improvements may be made in this direction.

SECTION (2)

Static Tests

Before treating the wiping, recording, and reproducing systems in further detail, it may be well to describe certain of the static tests that have been carried out. The term " static " is here applied to tests carried out upon either moving or stationary specimens of tape which do not involve the rate of change of the magnetizing current.

To show the magnetic properties of the tape, tests were made upon ring-shaped specimens. For example, a ring formed by winding 125 turns of tape to a mean diameter of 5 cm. has been used. The thickness of the ring was approximately 1 cm., and as the width of the tape is 0·3 cm. the cross-sectional area was 0·3 cm^2. The magnetizing winding consisted of 300 turns wound toroidally, and a search winding of 50 turns was connected to a fluxmeter.

The core was subjected to a sequence of magnetic fields such as might, in an ideal case, be experienced by a length of tape during the recording process. By observing the readings of the fluxmeter a family of curves was plotted showing the changes in the magnetic state of the tape over the whole range of the $(B - H)$ characteristic (Fig. 6).

The remanent induction retained after any given demagnetizing current has been applied to the saturated tape may be plotted against the field due to this current. Fig. 7 shows such a remanence curve (A) which can be deduced from the results shown in Fig. 6. It is of significance since it indicates the relation, in the static case, between the recording field and the remanent flux in the tape.

It might be thought that very different results would be obtained if tests were carried out upon a length of tape under normal conditions. This point has been examined by threading three coils on a single length of tape running on the machine, the first carrying a steady current adequate to bring the tape to saturation and the second being used to demagnetize it to any required degree. The circuit associated with this coil included a resistance, a switch, and a d.c. source. The third coil was used as a search coil and was connected to a ballistic galvanometer which was calibrated against a fluxmeter. The tape was run through the three coils at the normal speed of 1·5 metres per sec., the demagnetizing coil being unexcited. The switch completing the circuit of this coil was then closed and the change in remanent induction was calculated from the swing of the galvanometer. This was repeated for various values of demagnetizing current. Curve B in Fig. 7 shows the relation between the change of the remanent induction and the demagnetizing field.

The two curves in Fig. 7 show a very reasonable agreement, in view of the fact that the two specimens of the tape used were not necessarily identical in properties. Further, the calculation of the value of the field was only approximate, whilst in computing the values of the induction (B) it was assumed in both cases that the flux was uniformly distributed throughout the cross-section.

The effect of using a pole-piece instead of a coil for saturating the tape has been examined, using the method previously described to obtain the remanence curves shown in Fig. 8. These indicate the comparative effects of (A) A coil wound round the tape; (B) The same coil wound round a stalloy pole-piece with a flat tip; (C) The same coil wound round a pole-piece of which the thickness of the tip was reduced to about one-sixth of the total thickness.

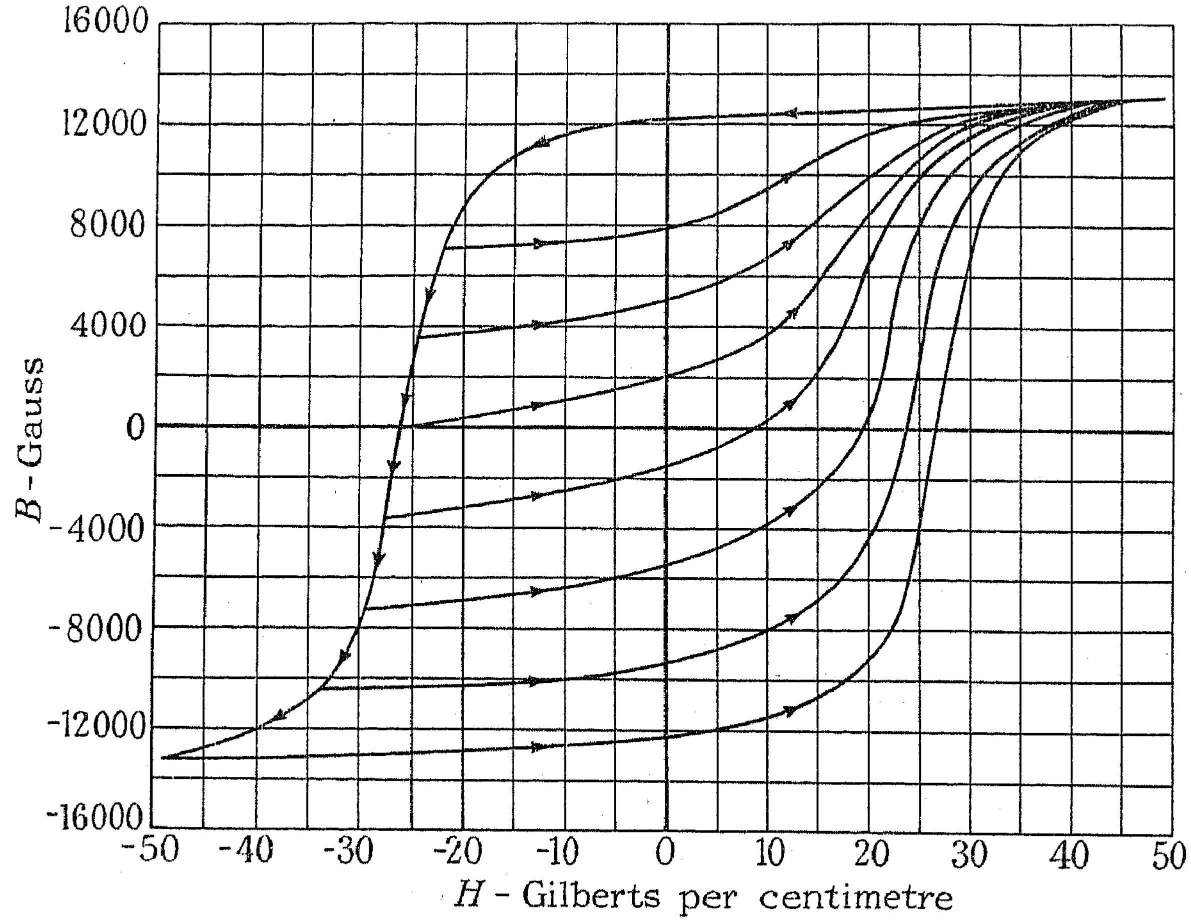

Fig. 6.—(B — H) characteristic for ring of tape.

It will be noticed that in curve C the slope is considerably reduced at the higher values of current. This is probably due to the effects of saturation at the tip of the pole-piece.

The situation will be somewhat more complex when a two-pole head is used. It was thought at one time that whenever two pole-pieces of opposite polarity were used for wiping or recording, as shown at the top of Fig. 9, the direction of magnetization of the tape leaving the heads was that of the magnetic flux between the two pole-pieces. It was assumed that since the stray flux on the running-off side of the heads must be lower in value than the flux between the pole-pieces, the former had little effect upon the final magnetic state of the tape. This, however, is only the case for a very small longitudinal gap between the pole-pieces.

Consider the case of a tape which has been fully magnetized by a field h_1 and is then passed through a two-pole wiping head in which the pole-pieces are widely staggered and of opposite polarity. Fig. 9 also shows the directions of the various fields to which the tape is subjected. As it approaches the leading pole-piece (P_1) it is subjected to the stray field h_2; then, between the pole-pieces, to the field h_3; and finally, as it leaves the head, to the stray field h_4 from the lagging pole-piece (P_2).

A curve (A in Fig. 9) has been plotted showing the relation between the current in the head and the resulting change of remanent induction in the tape ($B_{r\,max.} - B_r$) for positive and negative values of current.

For positive values of current the direction of h_3 is the same as that of the original saturating field h_1. The effect of h_2 is always largely counteracted by that of the

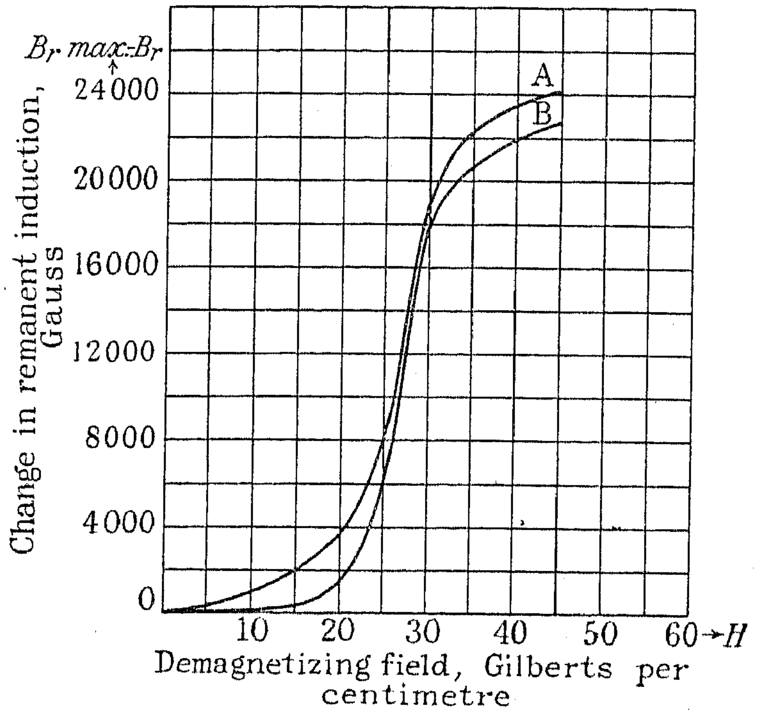

Fig. 7.—Comparison of remanence curves for ring and running tape.

A.—Ring of tape.
B.—Running tape.

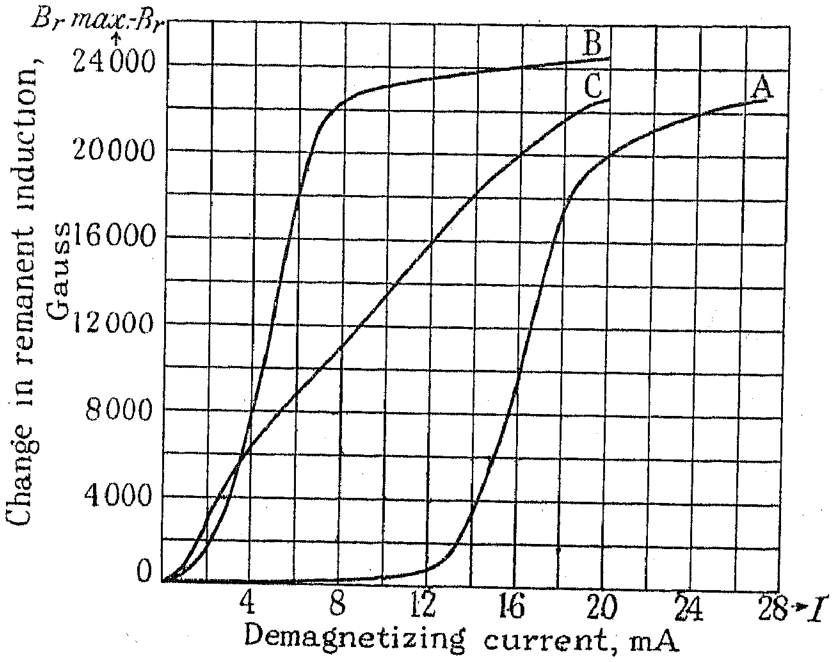

Fig. 8.—Comparison of remanence curves for coil and pole-pieces.

A.—Coil wound round tape.
B.—Same coil wound round saturating pole-piece on tape.
C.—Same coil wound round sharpened pole-piece on tape.

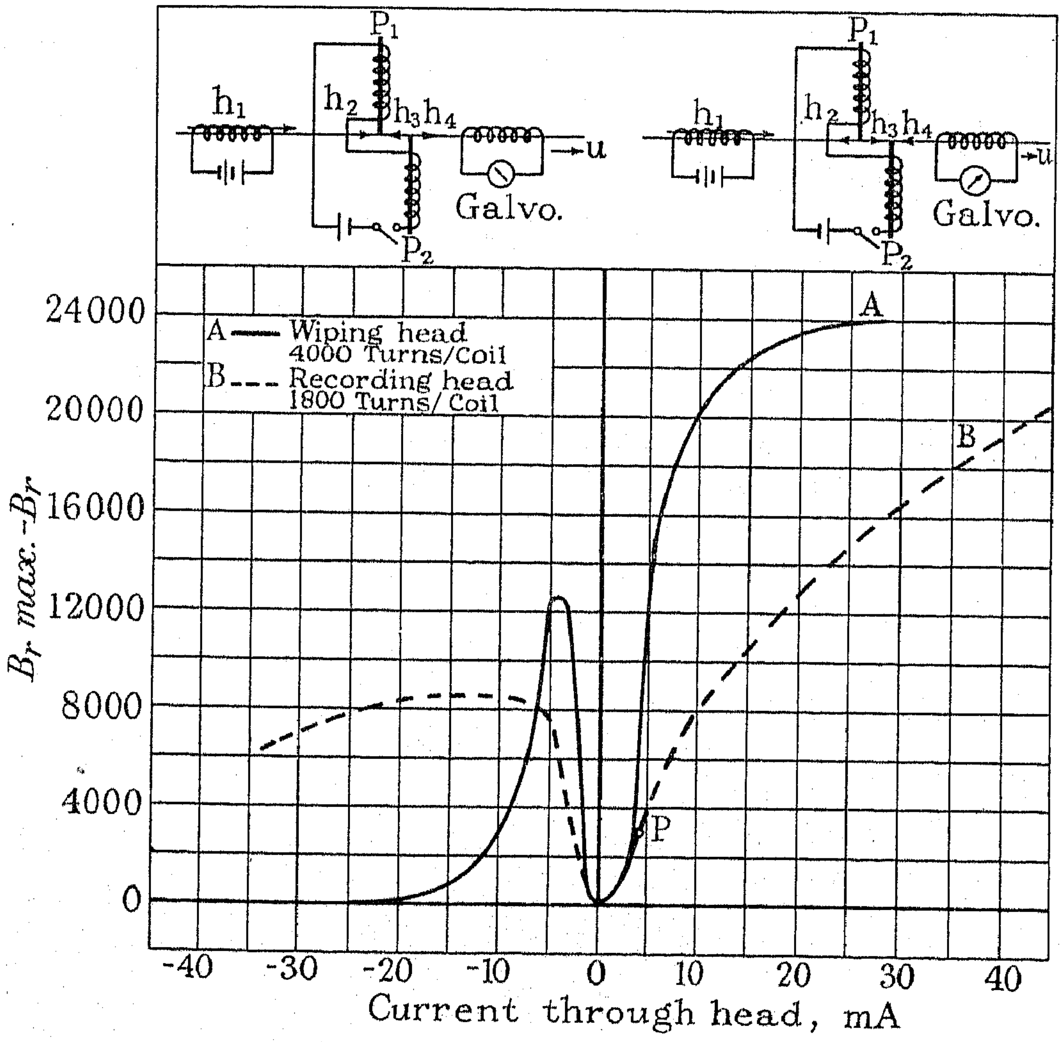

Fig. 9.—Remanence curves for two pole-pieces.

stronger field h_3, and the main change of remanent induction is produced by h_4. A value of current is reached at which the tape leaves the head fully magnetized in the direction of the field h_4.

For negative values of current, the direction of h_3 is opposed to that of h_1, h_2, and h_4. As the field h_3 is considerably greater than h_4 the rate of change of remanent flux with current is at first greater, and the curve rises more sharply, than in the previous case when h_4 was producing the change. A point is reached, however, when the tape tends to become saturated by h_3, so that further increases of current have less and less effect upon its state between the pole-pieces. The effect of h_4, however, still increases with an increase of current, and the curve bends over and returns gradually to the zero line as h_4 eventually becomes great enough to saturate the tape in the original direction. It is thus seen that, provided the saturating current is large enough, the final direction of magnetization of the tape is determined by the direction of the field h_4 and not by that of h_3.

A similar argument applies to the performance of a two-pole recording head with sharpened pole-pieces widely staggered. The curve for this head is shown as B in Fig. 9. From this it may be seen that when h_4 is in the same direction as h_1 the curve for a wiping head (A in Fig. 9) bends over and approaches the datum line, but with a recording head (B in Fig. 9) this tendency is very much reduced. The number of turns on each coil is 4 000 in the case of the wiping head and 1 800 in the case of the recording head. Apart from this, the difference in the behaviour of the two heads is probably due to saturation effects in the sharpened tips of the recording pole-pieces.

In practice, when using this double-pole system the value of demagnetizing current is such that when the a.c. recording current is zero the state of the tape leaving the recording head is represented by point P (Fig. 9).

The foregoing remarks immediately indicate the possibilities of various other arrangements of wiping and recording heads, including the use of one instead of two pole-pieces. A number of these arrangements have been tried with varying degrees of success, and will now be considered.

Wiping Systems

The purpose of the wiping head is to remove all previous flux variations in the tape, and to ensure that it is fully and uniformly magnetized in a longitudinal direction when it approaches the recording head.

Three alternative wiping systems are shown in Fig. 10.

(a) Coil.

In the centre of the coil the field will be almost entirely longitudinal, and the tape should be entirely saturated, provided that the field strength is adequate. In practice this method is not readily applicable, owing to the fact that the tape must be threaded through the coil.

(b) Single pole.

The distribution of the field from a single pole in contact with the tape is such that the tape is saturated in one direction as it approaches the pole-piece and in the opposite direction as it leaves it.

(c) Two poles with a gap between them in the direction of the motion of the tape.

In this case the tape experiences a further reversal in the direction of its longitudinal magnetization.

Both the single- and the two-pole systems can be made to give satisfactory results.

Recording Systems

A number of alternative arrangements and methods of excitation of pole-pieces are shown in Fig. 11, to which reference is made in the following description of their operation. These are grouped under three headings:—

(1) Single pole.
(2) Two poles with a wide gap between them.
(3) Two poles with a narrow gap between them.

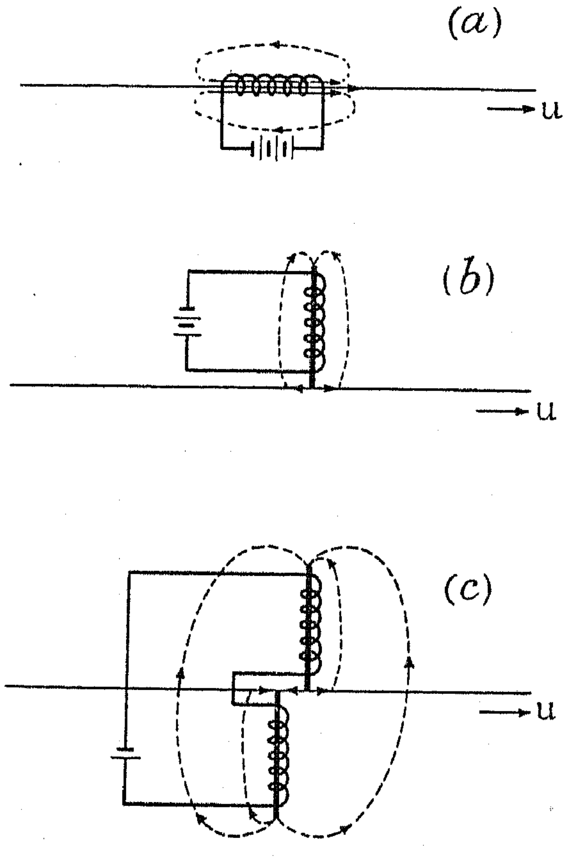

Fig. 10.—Wiping systems.

(*a*) Coil.
(*b*) Single-pole.
(*c*) Two-pole.

In group (1) the recording is carried out by the demagnetizing field on the trailing side of the pole-piece.

In group (2) it is largely carried out by the field on the trailing side of the second pole-piece.

In group (3) it is largely carried out by the main field between the pole-pieces.

The exact nature of the field distribution with any given arrangement is uncertain, but it would seem probable that in groups (1) and (2) the recording field rises sharply to a maximum near the trailing edge of the trailing pole-piece and then falls off rapidly. With a narrow gap, as used in group (3), the field distribution is much more nearly square-topped.

In the case of group (3), with the exception of system No. (3*a*), any lengths of tape demagnetized by the main recording flux between the pole-pieces are subsequently subjected to the stray field on the trailing side of the second pole-piece, which is, of course, in the opposite direction. If the pole-pieces are accurately ground and adjusted, however, this stray field should be exceedingly small when the gap is very narrow, but otherwise it may well become appreciable.

***Group* (1).**

The simplest recording head consists of a single pole with a single magnetizing winding in which the alternating current is superimposed upon the direct current. If wiping is carried out with a single north pole, the recording pole must be a south pole, and vice versa. As the tape approaches the recording pole-piece it is subjected to the field h_3 which is in the same direction as the saturating field h_1, so that when it reaches the centre of the pole-piece, where the longitudinal field is zero, the tape will still have approximately its maximum possible remanent longitudinal magnetization. As it moves away from the pole-piece it is subjected to h_4, which is in the demagnetizing direction and carries out the recording.

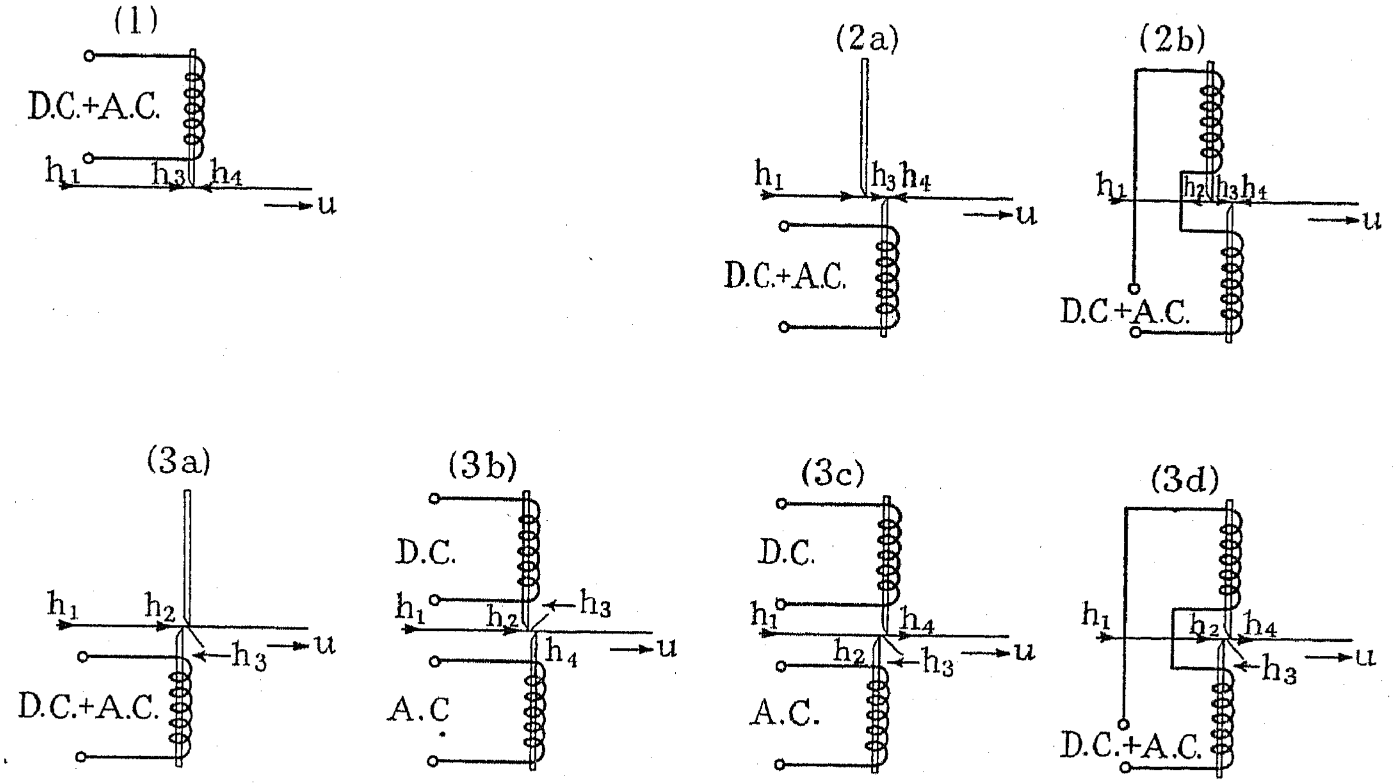

Fig. 11.—Recording systems.

***Group* 2).**

With both these arrangements the exact setting of the longitudinal gap between the pole-pieces is not of fundamental importance, and about 3 mm. appears to give the best results. The function of the first pole-piece is not very clear, but results appear to be improved by its presence. In all probability it serves mainly to reduce the spread of the recording field h_4.

In case (*a*) the operation is almost identical with that of the single pole-piece of group (1).

In case (*b*), which is known as the " D.P." system, the tape approaching the first pole-piece is demagnetized by h_2, but as it passes under the pole-piece it is subjected to h_3, which is in the saturating direction. The distribution of this field is such that it is at a maximum under the trailing edge of the first pole-piece and under the leading edge of the second pole-piece, while in between it may be slightly lower. Since the thickness of the pole-piece is small compared with the shortest wavelength to be recorded, the value of the alternating components of h_2 and h_3 will not have appreciably changed as the tape passes under the first pole-piece, and the demagnetization due to h_2 is counterbalanced by the saturation due to h_3. Recording is then carried out by the second pole-piece in much the same manner as with a single pole-piece.

With the arrangement of group (2) it seems that slightly better results are possible at the higher frequencies than with group (1), but at lower frequencies the converse holds.

***Group* (3).**

(*a*) This system has been termed the single-pole and idle-pole, or " S.P.I." system. The second pole-piece is unexcited and therefore does not produce any net reversed stray field, but serves to concentrate the recording field h_3.

(*b*) and (*c*) The advantage of these two systems is that the d.c. and a.c. components of demagnetizing current are separated, so that there will be a reduction in the asymmetrical distortion of the waveform due to the saturation of the tip of the a.c. energized pole-piece.

In (3*b*) the final stray field to which the recorded tape is subjected is alternating, while in (3*c*) it is approximately constant in magnitude and direction. On the other hand, considering the effect of the stray field h_2 as the tape enters the head, in the case of (3*b*) this will be constant and in the same direction as the original saturating field h_1, so that the recording is carried out upon tape that is always in the same magnetic state. In (3*c*) h_2 is alternating, and this condition does not apply, although in any case the stray fields should be very low.

(3*d*). In this arrangement the method of excitation is similar to that of (2) (*b*) but has reversed polarity.

The tape recorded by h_3 is subjected to the field h_4 which has both d.c. and a.c. components, each of which is liable to be lower than the corresponding component in (3*b*) and (3*c*).

In group (3) the adjustment of the head is very critical and the procedure in each case is to set the poles opposite to one another and then to increase the longitudinal gap between them until the output from the reproducing head at 6 000 cycles is a maximum.

The variation in the results obtained with the four arrangements of group (3) is not very great. It would appear, however, that with all of them a higher level may be recorded than with either group (1) or group (2), but that under practical conditions distortion at low levels is liable to be slightly worse.

Reproducing Systems

Three types of reproducing systems are mentioned and are illustrated in Fig. 12.

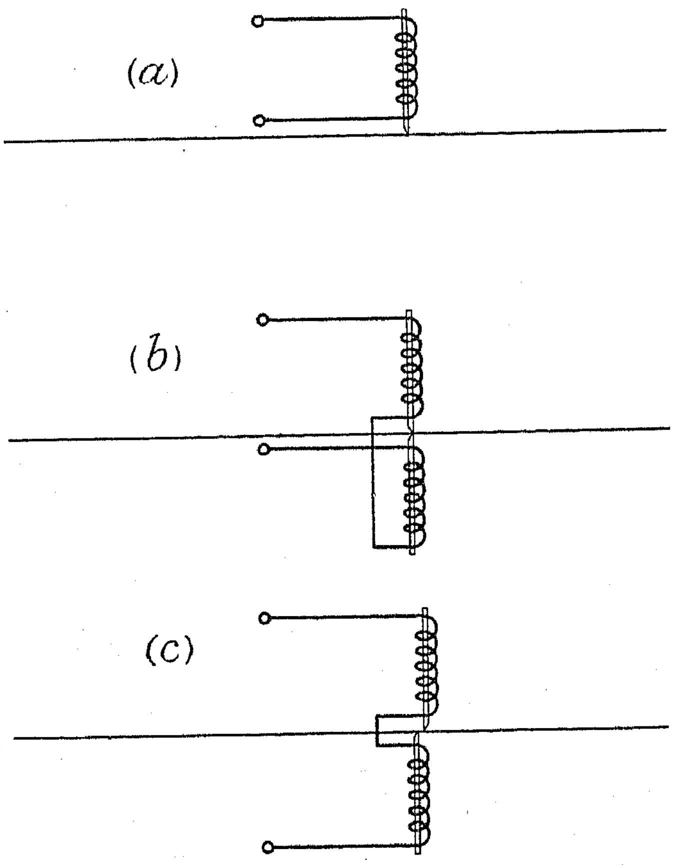

Fig. 12.—Reproducing systems.

(i) *Single pole-piece.*

The operation of this has already been briefly described in Section (1). This system has the very great advantage that extremely consistent results may be obtained, since no adjustments have to be made.

(ii) *Two pole-pieces opposite each other.*

The operation of this system is, in theory, almost identical with that of a single pole-piece. The coils round the pole-pieces are connected in series opposing, so that the e.m.f.'s generated are additive. It would appear from practical experience that the presence of pole-pieces on either side of the tape results in a higher efficiency of pick-up at high frequencies than is the case with a single pole-piece. The system has the disadvantage that the pole-pieces must be very accurately set or consistent results will not be obtained.

(iii) *Two pole-pieces with a gap between them.*

In this case the coils are connected in series aiding. The net e.m.f. induced in the two coils is proportional to the difference between the rates of change of the number of flux-turns in each coil. This difference will be a maximum when the distance between the pole-pieces is adjusted to be equal to one-half of the wavelength of the flux alternations in the tape. Although this system was used for some time, it is very difficult to adjust the distance between the pole-pieces very accurately, so that a consistently uniform frequency characteristic is not always obtained from the system.

In addition to the simple forms of heads that have been described above, there are a large number of other possibilities. These include ring-type heads, which consist of a ring-shaped core carrying a single winding and having a small gap in the centre of the portion in contact with the tape. The plane in which the ring lies is parallel to the length and thickness of the tape. This type of head has been developed in Germany in connection with the use of magnetic dust-coated film instead of steel tape.

After experimental work in this country with different types of heads, two recording systems, known as the "D.P." and "S.P.I." (2*b* and 3*a*, Fig. 11) have emerged; these are used by the British Broadcasting Corporation and the Marconi Co. respectively. It is fully realized that both these systems are subject to serious disadvantages from both a theoretical and practical point of view, but at the time of writing the authors have found no other system giving more satisfactory results under service conditions.

In each case a single-pole-piece reproducing head is used.

The Complete Recording Chain

Individual systems of wiping, recording, and reproducing, have already been discussed. The problem of obtaining the best results from the complete recording and reproducing chain for any given arrangement of the various heads involves the question of the adjustment of the recording head, and the working point on the ($B - H$) curve, as well as the best possible recording frequency characteristic and the corresponding reproducing frequency characteristic.

It has already been pointed out that the total volume range of any system of recording is bounded on the one hand by amplitude non-linearity and on the other hand by considerations of background noise. To a certain extent the frequency range will also be bounded by the same two factors. For although, within limits, it is possible to apply frequency correction circuits to the reproducing amplifier which will give a frequency characteristic that is substantially flat over a very wide range, in so doing the signal/noise ratio may be reduced.

The frequency characteristic of the recording amplifier should be chosen to allow the recording level to be as high as possible for a given amount of distortion under all programme conditions, so that the maximum signal/noise ratio is obtained. This condition involves not only the characteristics of the recording chain but also the energy distribution of the programmes to be recorded, which will be very varied. Numerous experiments have been carried out in this connection, and it is generally accepted that the maximum energy in the frequency spectrum of speech and music occurs in the neighbourhood of 300 cycles per sec.

Curve A of Fig. 13 shows the variation with frequency of the peak energy levels occurring in a typical orchestral programme as determined by Harvey Fletcher,* and curve B shows the maximum level, consistent with linearity, which may be applied to the input of the recording amplifier for any single frequency, deduced from the results of experiments on the D.P. system. It must be noted, however, that this curve does not show the modulation effects of one frequency upon another, as will be explained later.

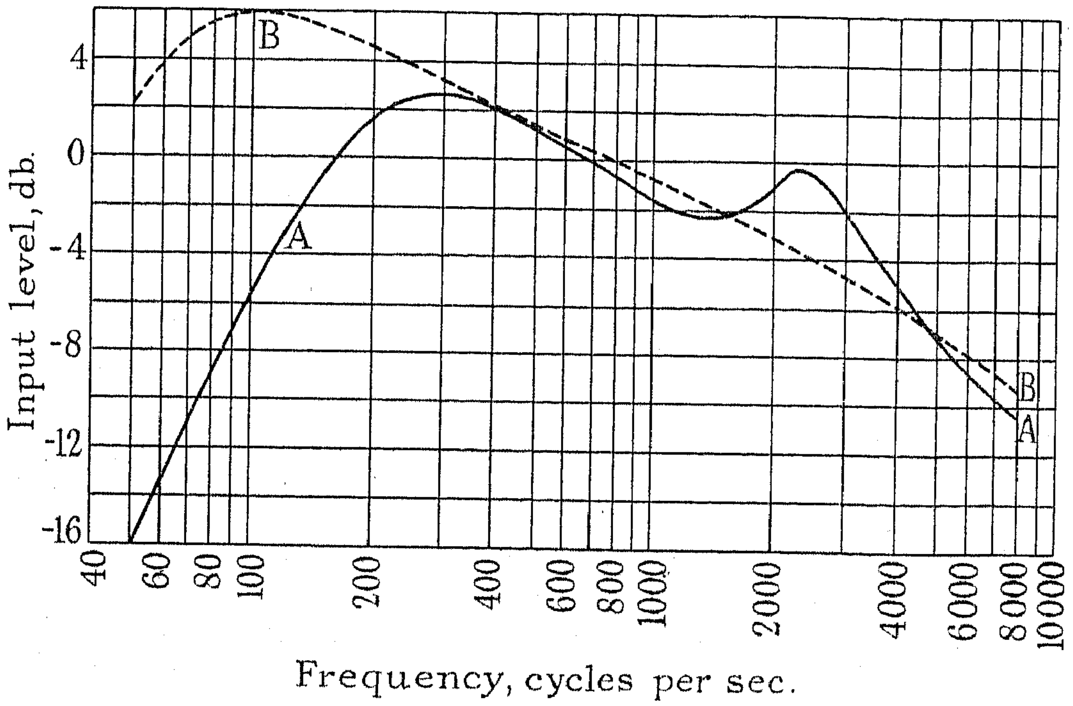

Fig. 13.—Peak energy levels occurring in orchestral programme.

In carrying out comparative tests upon recording heads, measurements of the reproduced signal must be made, since a satisfactory direct method of measuring the magnetic state of the tape has not been developed. The frequency discrimination of the reproducing system must be taken into account, however, when measuring the relative levels of different frequencies reproduced, but this is, of course, dependent upon the recording characteristic if the overall frequency response is to be flat. To avoid this complication a form of test is often required which involves a single frequency only. In this connection the linearity curve has proved most useful. If the input level to the recording head is plotted against the output level from the reproducing head to the same scale, the resultant curve should be a straight line at 45° to either axis. Departure from this line in either direction indicates amplitude non-linearity which would give rise to distortion. With this system of recording it is usual for the reproduced level to rise sensibly linearly with the recording level up to a point beyond which it falls below the linear value. Eventually a maximum reproduced level is reached such that further increases of recording level produce a reduction in the reproduced level.

These linearity curves can be used first of all to determine the most suitable working point on the $(B - H)$ curve, as well as the best adjustment of the recording head. An examination upon a cathode-ray oscillograph of the changes in the reproduced waveform while making these adjustments is often of great value.

The authors have found, however, that the results of listening tests when recording programmes do not always agree with the results of tests using single frequencies. In fact, a single-tone test is inadequate to determine the distortion of the particular system under consideration, and experiments using two tones should be made. This statement may be made clearer by citing an example. In Fig. 14 input/output linearity curves are shown for three frequencies with different values of demagnetizing current, using the D.P. recording system. It will be seen that the output level is dependent upon the value of this current throughout the straight and useful portion of the curves. Bearing in mind the process of recording described earlier in the paper, imagine that two tones having frequencies of, say, 100 and 4 000 cycles per sec. are being recorded simultaneously. The variation in the value of the recording current due to the low frequency effectively alters the mean value of the demagnetizing current upon which the high-frequency alternations are superimposed. This tone is thus modulated during the recording process at 100 cycles per sec., which produces a type of distortion that is not directly shown up by experiments using a single tone.

The recording characteristic used in practice will also depend, to a certain extent, upon the method used for monitoring the input level to the recording amplifier. Since it is possible to listen to the reproduced signal a fraction of a second after recording, it is usually convenient to make use of this facility to determine the highest permissible recording level for control purposes. In order to do this, overloading must occur before suitable adjustment can be made. It has been found that with aural control the quality of the reproduction, as a whole, suffers least if overloading is first apparent at the lower frequencies, rather than simultaneously throughout the frequency range. Under these circumstances the recording characteristic may be modified at the lower frequencies to produce this effect.

---

* "Some Physical Characteristics of Speech and Music," *Bell System Technical Journal*, 1931, vol. 10, p. 349.

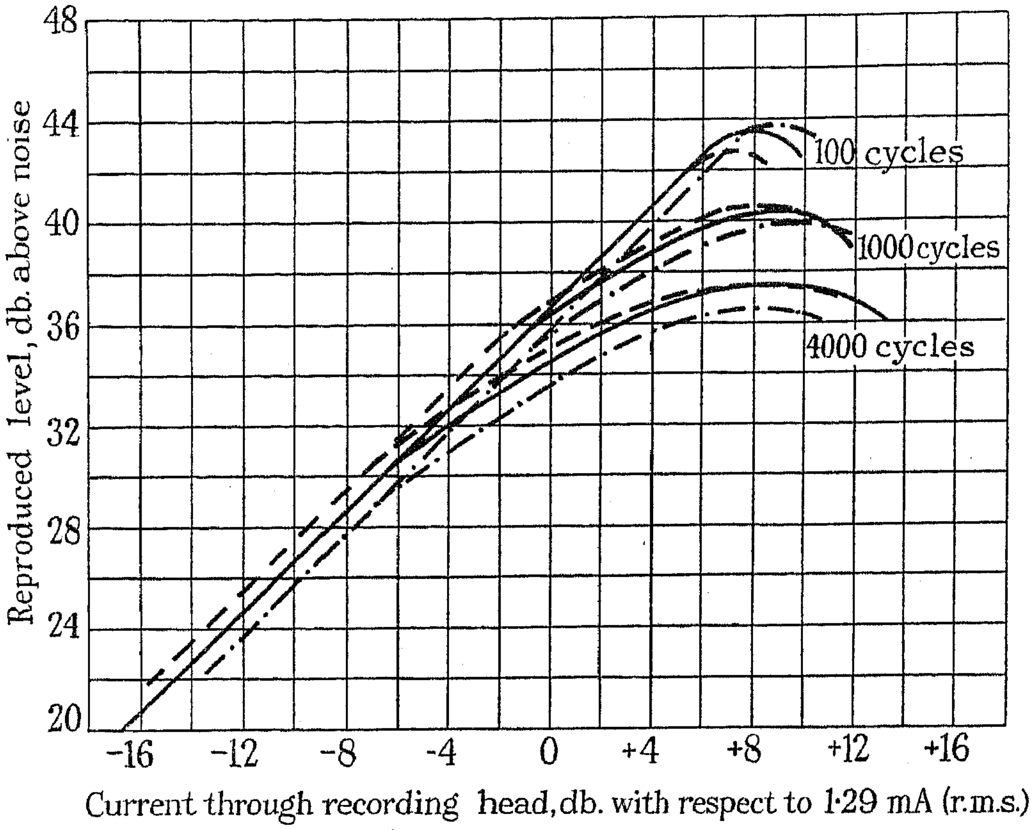

Fig. 14.—Input-output linearity curves for D.P. system.

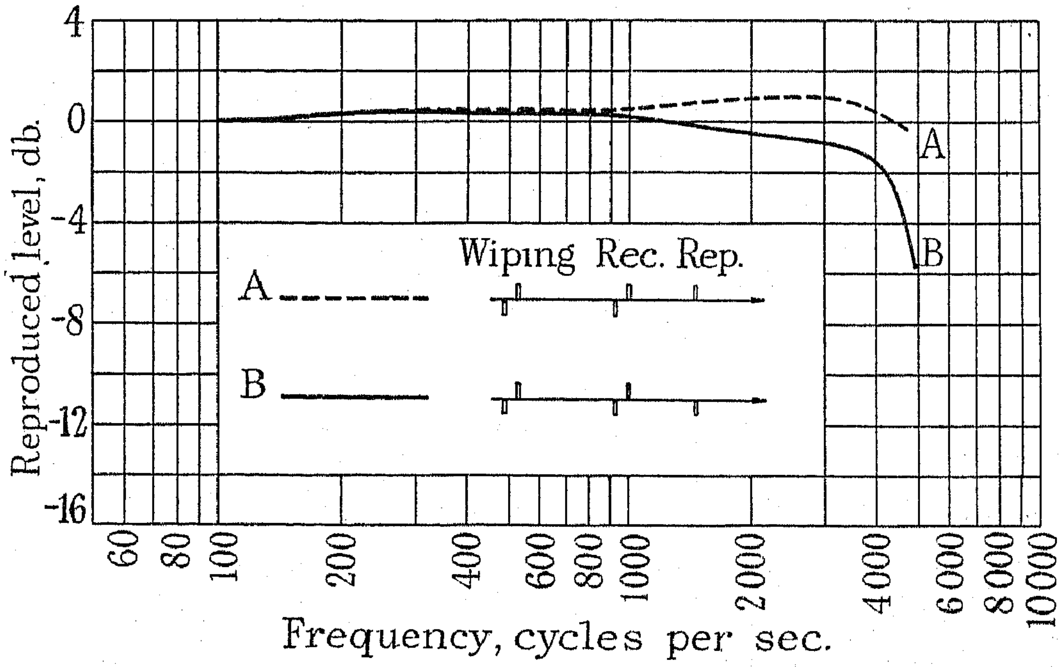

Fig. 15.—Effect of asymmetrical recorded flux distribution.

The frequency characteristic of the reproducing amplifier must be adjusted to give an overall frequency response that is sensibly flat throughout the required range, while maintaining a satisfactory signal/noise ratio.

When the reproducing frequency characteristic is such that the reproduced noise is most noticeable at the higher frequencies, it is sometimes found that by attenuating all frequencies above, say, 5 000 cycles per sec., the improvement in the signal/noise ratio more than compensates for the reduction of frequency range. This has actually been done in the case of the D.P. reproducing amplifier, of which the frequency/response characteristic is given in Fig. 16.

Various quantitative measurements may be carried out upon the complete recording and reproducing chain, of which the most instructive are probably the input/output linearity curves, and two tone tests for a high and a low frequency. It is thought that the more usual harmonic tests are of secondary importance, since their main object is to show up non-linearity which will result in the production of intermodulation tones.

The ultimate test of any recording system, however, is to listen to the reproduction of a number of recorded programmes and to compare them with the direct programmes under identical listening conditions.

Some of the details and some of the results of tests upon the D.P. and S.P.I. systems now in use are given below.

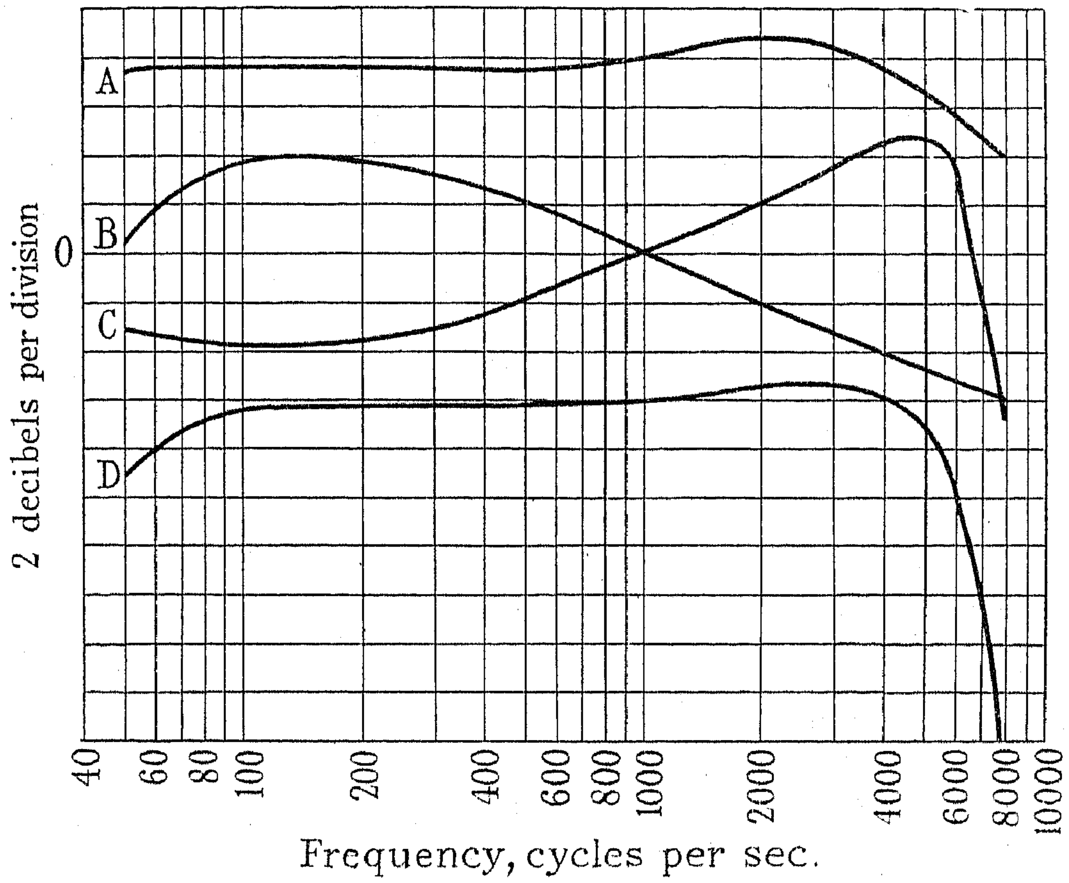

Fig. 16.—Frequency/response characteristic for D.P. system.

The D.P. and S.P.I. Systems of Recording.

These two systems are illustrated in Figs. 11, 2*b* and 3*a*, and a brief explanation has already been given of their operation. In practice they are both used with a single-pole reproducing system. The reproducing pole-piece runs on the same side of the tape as the lagging recording pole-piece in the D.P. system, and the energized pole-piece in the S.P.I. system. The effect of placing the reproducing pole-piece on the other side of the tape is shown in Fig. 15. Curve A shows the overall frequency/response characteristic of the D.P. system taken with the reproducing pole-piece on the same side of the tape as the lagging pole-piece in the recording head. Curve B shows an equivalent curve, the reproducing pole-piece being on the opposite side of the tape. It will be seen that there is a considerable loss at the high frequencies in the latter case, but that the reproduced level at the lower frequencies is the same. This suggests that where the wavelength is great compared with the thickness of the tape the external flux distribution is symmetrical on either side, but that where the wavelength is comparable with the thickness the density is greater on the side of the tape that carries the lagging pole-piece. Therefore, in practice, the reproducing pole-piece is always used on this side.

It has been found that a slight improvement in results is obtained when the recording and reproducing pole-pieces are reduced at the tips in width as well as in thickness, so that the whole of the width of the tape is not uniformly demagnetized by the recording head.

Stalloy is used for recording, and permalloy for reproducing, pole-pieces, which are stamped from 0·015-in. sheet (i.e. 0·37 mm.) and are 3 mm. wide and 19 mm. long. The tips are reduced in width to 2 mm., and in thickness, i.e. in the direction of motion of the tape, to 0·06 mm. They are, of course, heat-treated after shaping. The best value for the demagnetizing current has been found to be 4 mA for the D.P. system and 3 mA for the S.P.I.

The frequency/response characteristics of the two systems are shown in Figs. 16 and 17. In each case the four curves refer to the current in the recording heads (A), the magnetic systems which include the recording heads, the tape and the reproducing heads (B), the reproducing amplifiers (C), and the overall chains (D), and are drawn one below the other at entirely arbitrary intervals.

Measured under these conditions, input/output linearity curves for the D.P. system have already been given in Fig. 14, while Fig. 18 shows the results of two tone tests using 105 cycles and 4 000 cycles per sec., Curves A′, B′ and C′ show the root-sum-squared level of the intermodulation tones [r.s.s. (4 000 ± 105*n*)] for varying output of the low-frequency tone corresponding to levels of the high-frequency tone given by curves A, B, and C, respectively. It appears that the r.s.s. level of the intermodulation tones is proportional to the reproduced level of both the recording tones. With the S.P.I. system results of the same order are obtained, but the distortion is in general lower for very high levels of the higher frequency and higher for very low levels. In this case it is definitely advantageous to record with a recording frequency characteristic that rises at the higher frequencies, as may be seen in Fig. 17.

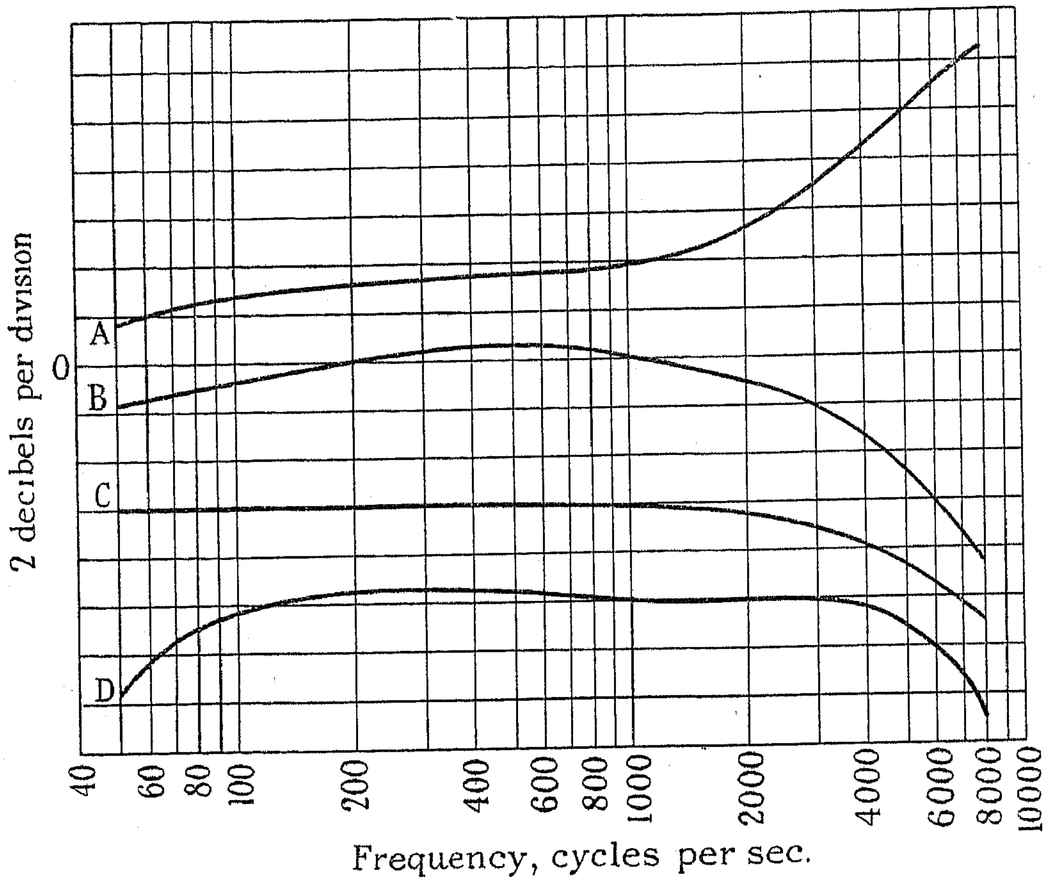

Fig. 17.—Frequency/response characteristics for S.P.I. system.

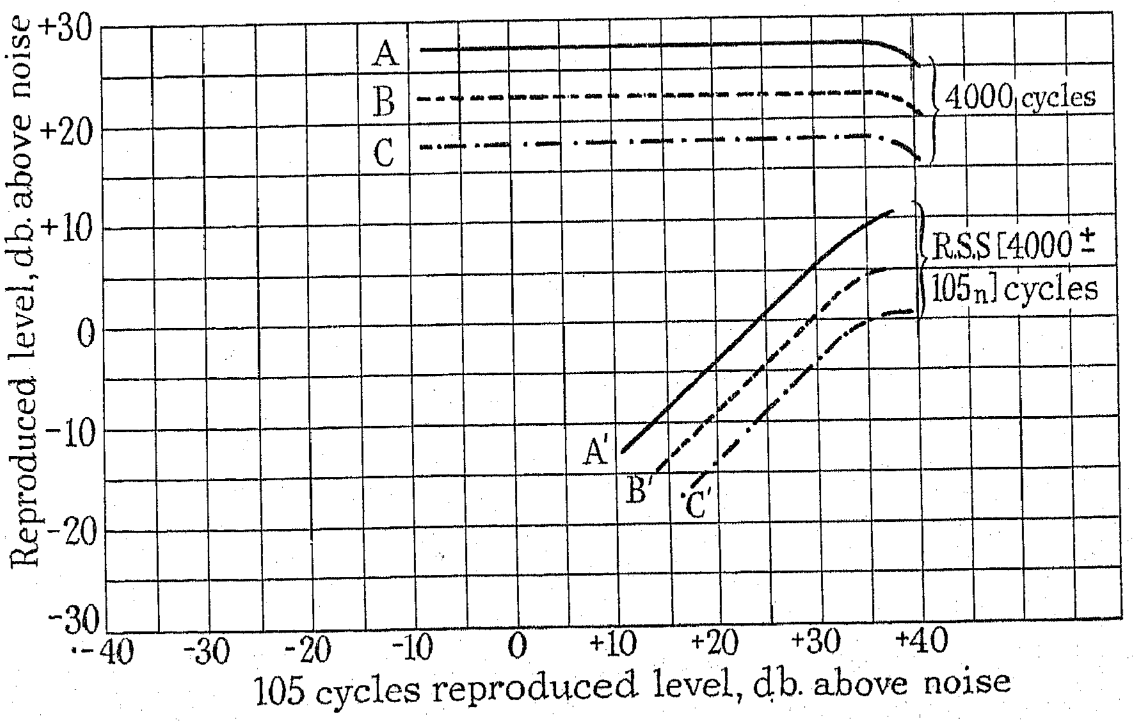

Fig. 18.—Levels of intermodulation tones reproduced when a tone of 105 cycles per sec. is recorded together with a constant level of 4 000-cycle tone on the D.P. recording system.

A.—Recording tone of 4 000-cycle tone:— 9 db./1·29 mA.
B.—Recording tone of 4 000-cycle tone:—14 db./1·29 mA.
C.—Recording tone of 4 000-cycle tone:—19 db./1·29 mA.

Measurements of the signal/noise ratios obtainable in practice with the two systems give figures ranging from 35 to 45 db.

In spite of the fact that from test-results on tone it would appear that a greater amplitude range can be obtained with the S.P.I. than with the D.P. system, it is found that under practical conditions there is very little to choose between them, and both systems are in use at the present moment.

Plate 1

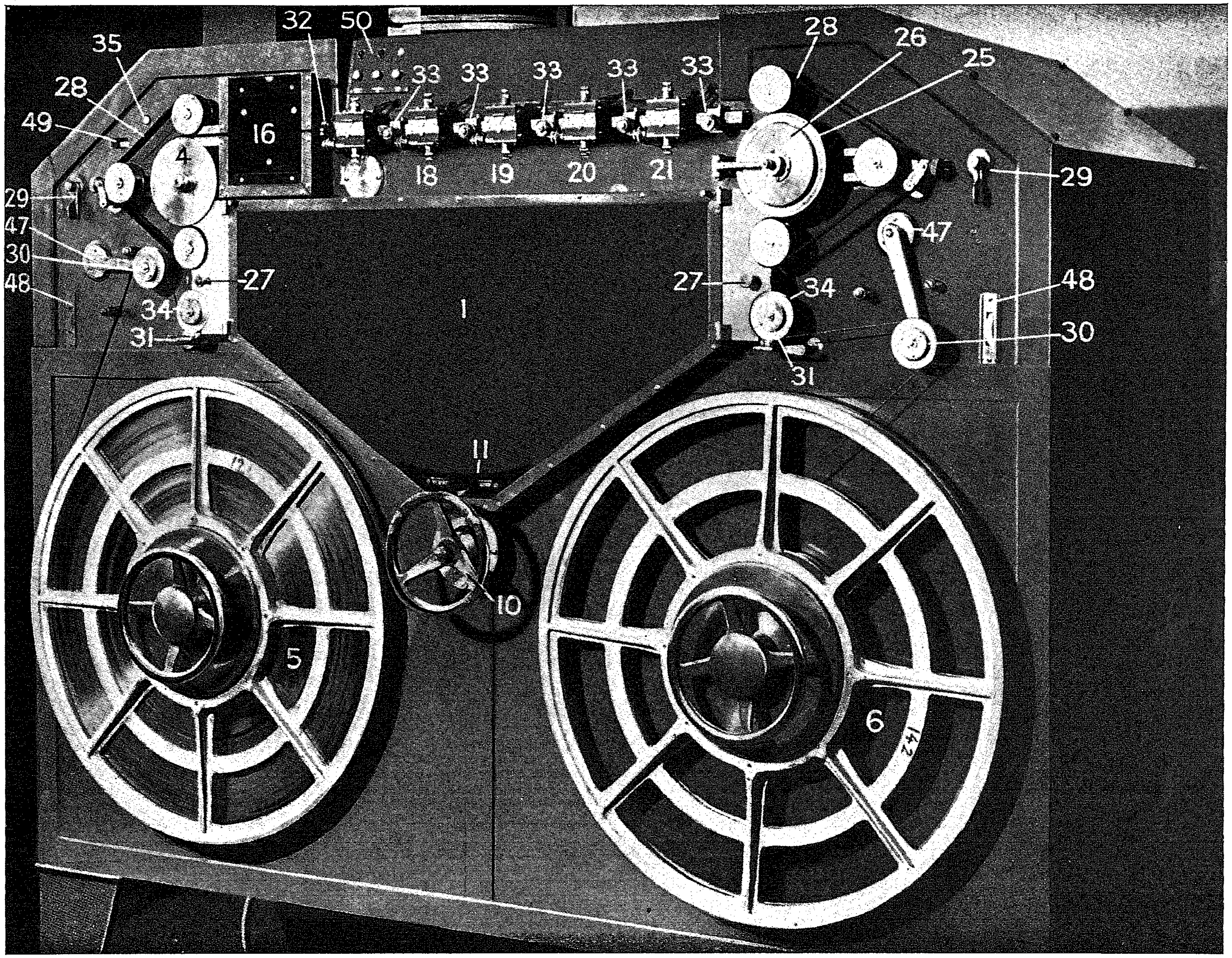

Fig. 19.—Front view of machine.

1. No. 1 tape reservoir.
4. No. 2 tape-drive wheel.
5. Left-hand tape reel.
6. Right-hand tape reel.
10. Starting handle.
11. Contact, No. 1 reservoir.
16. No. 2 tape reservoir.
17. Wiping head.
18 and 19. Recording heads.
20 and 21. Reproducing heads.
25. Tape drive wheel.
26. Tape drive wheel clutch.
27. No. 1 tape reservoir entry door.
28. Tape drive belt.
29. Quick-release toggle.
30. Spring jockey pulley.
31. Tape tension brake.
32. Friction pad.
33. Tape guide.
34. Lower guide wheel.
35. Pilot light.
47. Tape jockey-arm bearing.
48. Tape tension brake adjustment.
49. Tape travel indicator.
50. Cue lights.

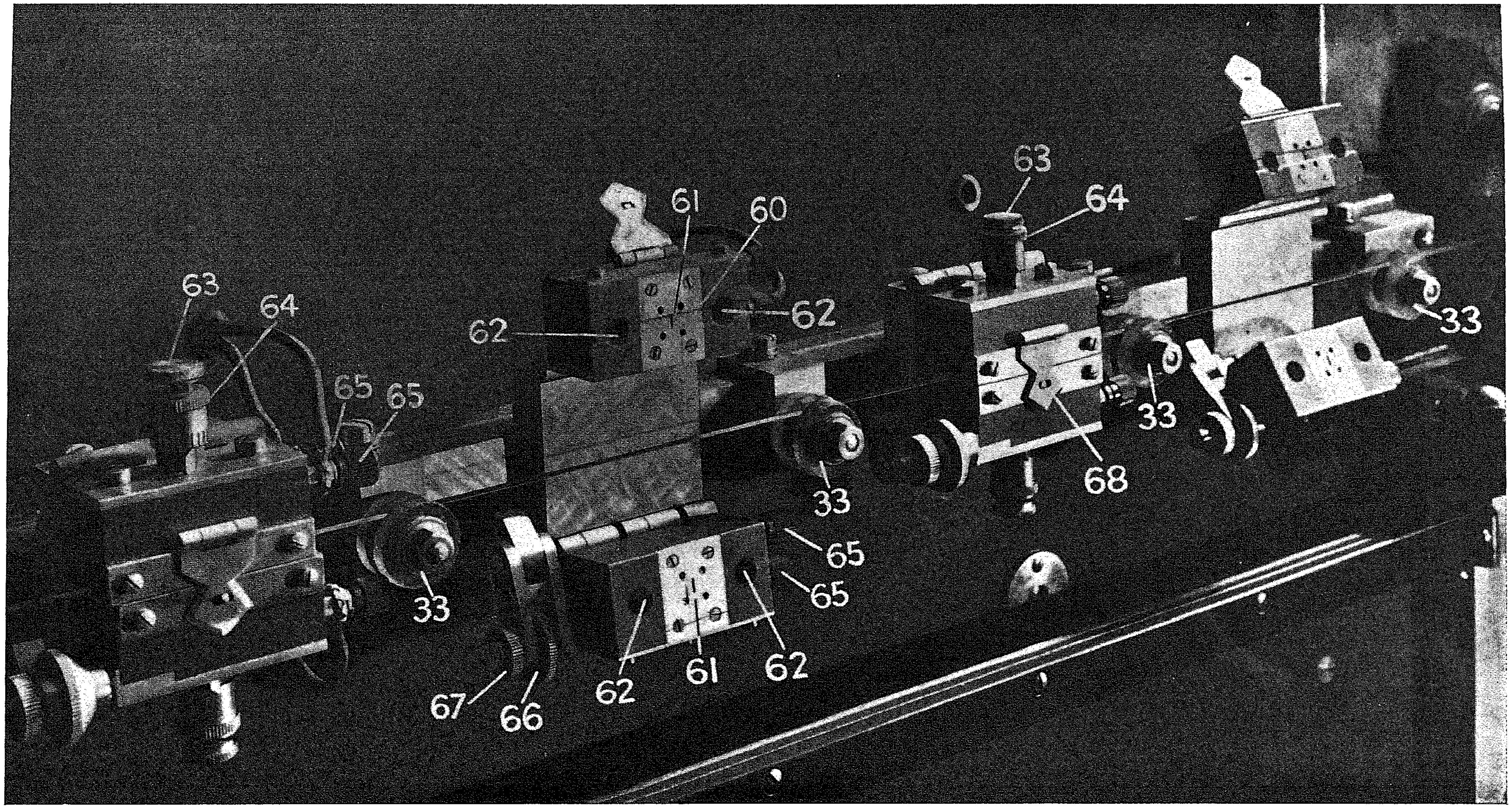

Fig. 20.—Recording and reproducing heads.

33. Tape guide.
60. Tape groove.
61. Slot for pole-piece.
62. Felt pad.
63. Pressure pin.
64. Spring-tension adjuster.
65. Terminals.
66. "Stagger" adjusting screw.
67. Lock nut.
68. Clamp.

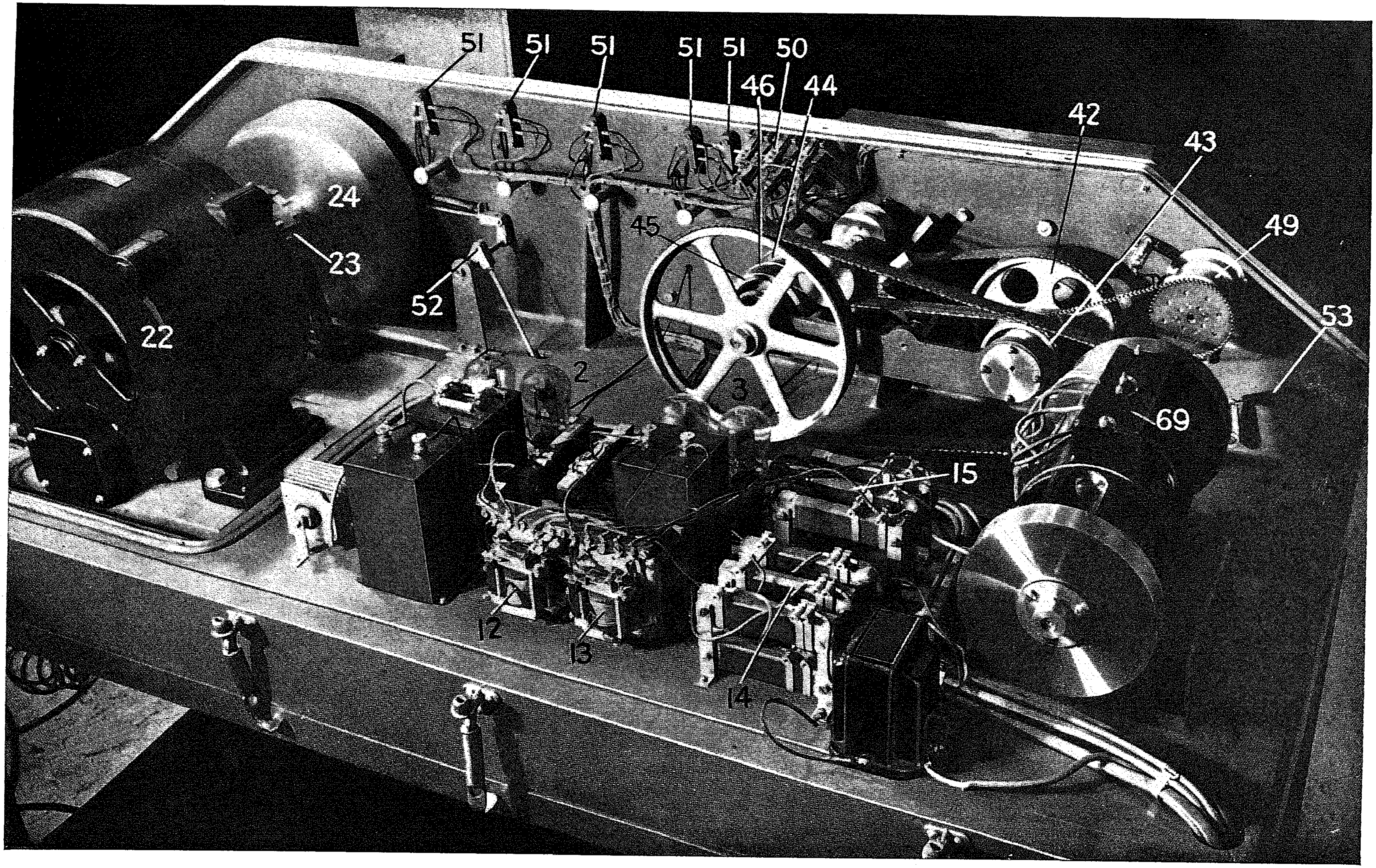

Fig. 21.—Back view. Upper deck (covers removed).

2. Thyratrons (No. 1 System).
3. Thyratrons (No. 2 System).
12. Relay, No. 1 System.
13. Relay, No. 2 System.
14. No. 1 motor resistances.
15. No. 2 motor resistances.
22. No. 3 motor (constant feed).
23. Flexible coupling.
24. Oil-damped flywheel.
42. Forward-driven pulley (No. 2 tape drive wheel).
43. Backward-driven pulley (No. 2 tape drive wheel).
44. Forward driving pulley (No. 2 tape drive wheel).
45. Backward driving pulley (No. 2 tape drive wheel).
46. No. 2 automatic dog clutch.
49. Tape travel indicator.
50. Cue lights.
51. Head connection blocks.
52. Tape drive wheel clutch-operating mechanism.
53. Quick-release toggle adjustment.
69. No. 2 motor.

Plate 4

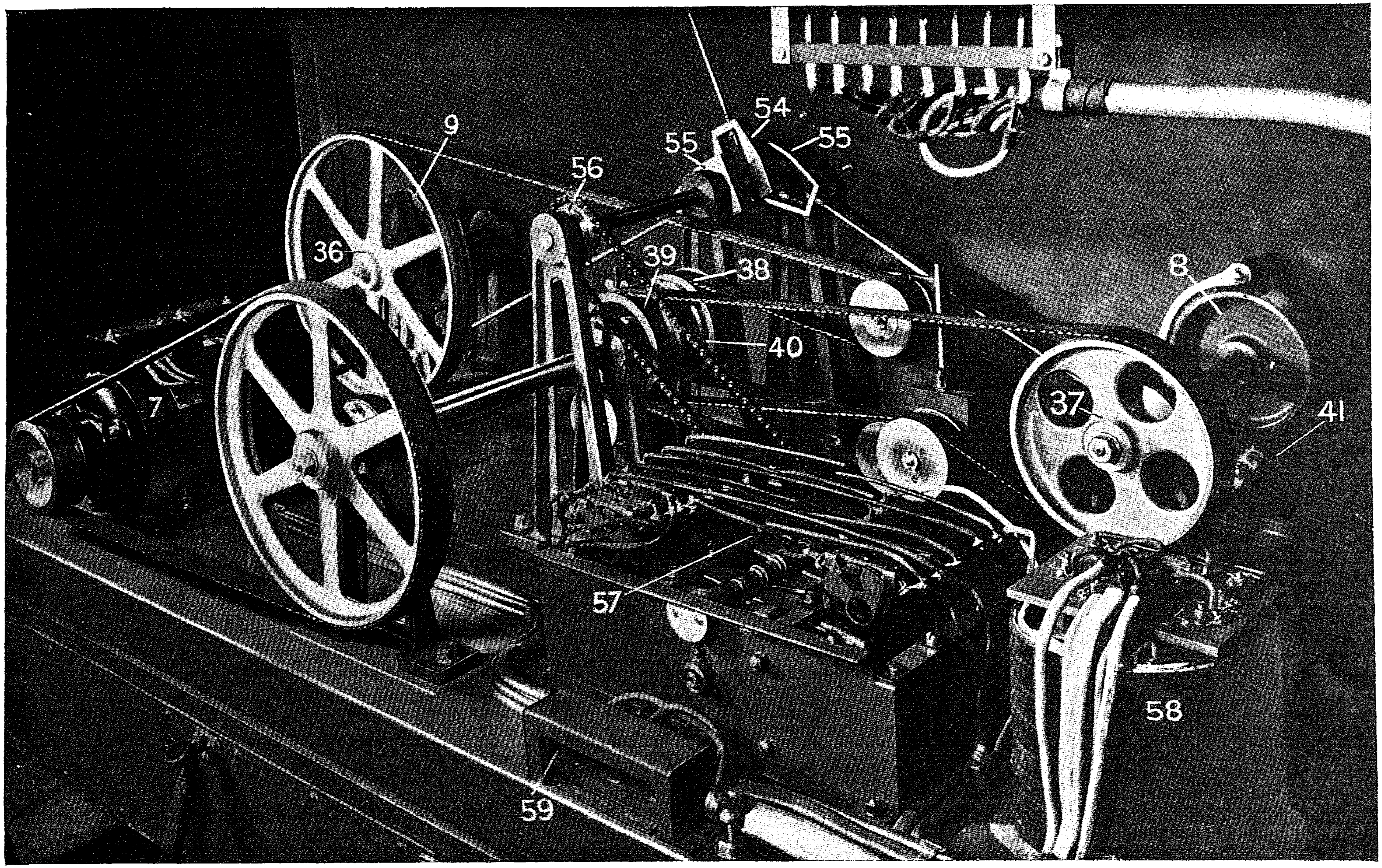

Fig. 22.—Back view. Lower deck (covers removed).

7. No. 1 motor (winding).
8. Brake, left-hand tape reel.
9. Brake, right-hand tape reel.
36. Right-hand tape reel driven pulley (forward).
37. Left-hand tape reel driven pulley (rewind).
38. Right-hand reel driving pulley (forward).
39. Left-hand tape reel driving pulley (rewind).
40. No. 1 automatic dog clutch.
41. Brake springs.
54. Cam-operated mechanism for tape drive wheel clutch.
55. Cam-operated mechanism for brakes.
56. Operating mechanism for "forward-reverse" switchgear
57. "Forward-reverse" switchgear.
58. Power transformer (thyratron circuits).
59. Main 3-phase connection block.

SECTION (3)

The Tape

The results of the recording system as a whole are to a great extent dependent upon the quality and uniformity of the steel tape.

A tungsten magnet steel is used, having the essential properties of high coercivity and remanence. It is melted in an acid open-hearth furnace from pig-iron produced in an electric charcoal blast-furnace from the purest ores, and scrap iron of known analysis. It is first of all hot-rolled under temperature control and then cold-rolled, cut, and polished, undergoing various heat treatments during the process of such a nature that the possibility of magnetic ageing is reduced to a minimum This allows a tape to be stored for a number of years after recording without noticeable deterioration of the quality of the recording. The final dimensions of the tape are 0·08 mm. thick and 3 mm. wide, and it is rolled into lengths of about 1 000 metres. Each spool of tape used for recording has to run for about half an hour at a speed of 90 metres a minute, so that some 2 700 metres are required. This means that even in a new tape there will be at least two joints, which are made by means of silver solder.

In the past, if breaks occurred in the tape during normal running, soft-soldered joints were made. A system of welding the tape which is now being tried out, however, gives most encouraging results.

Design of Recording Machine

The mechanical requirements to be fulfilled by the machine may be summarized as follows:—

(1) The tape must move through the heads at a constant velocity during both the recording and the reproducing processes.

(2) The tape must be unwound from, and rewound on to, the spools without undue tension being applied to it.

(3) A practically instantaneous start from rest is desirable in order to facilitate the commencement of a reproduction at a predetermined cue.

(4) A rewinding speed considerably in excess of the forward running speed should be available so that two machines can be run continuously in tandem if necessary.

Experience has shown that the foregoing requirements are best covered by the use of three separate mechanical drives, one for unwinding the tape from the trailing reel, another for maintaining the velocity of the tape constant through the heads, and the third for winding the tape on to the leading reel. The constant-speed system is virtually isolated from the unwinding and winding systems by the provision of reservoirs, one preceding it and one following it, and the speed of the winding and unwinding motors is automatically regulated so that both reservoirs always contain a loop of tape.

These two motors are of the induction type, with their rotor windings brought out to three slip-rings and their brushes connected to external resistances arranged as a three-phase star-connected system. With the full resistance in circuit the speed of the motor is such as to make the speed of the tape less than it is in the constant-speed drive system, but provision is made for short-circuiting part of the resistance. The winding and unwinding systems thus each have two forward speeds, one greater and one less than that of the constant-speed drive. By causing the motors to alternate between their two speeds a sufficient loop of tape is always maintained in both reservoirs. The switching in each case is carried out by a relay, the operation of which is controlled by a gas-filled thermionic valve of the thyratron type. This consists essentially of a half-wave rectifier, but by applying sufficient negative bias to its control grid the anode current can be completely suppressed. Such a bias is normally applied and, the relays being unenergized, the unwinding motor runs normally at its higher speed and the winding motor at its lower speed. Therefore loops form in both reservoirs. When the loop of steel tape becomes long enough, it touches a contact at the bottom of the reservoir, which is connected to the grids of the associated thyratrons, and short-circuits the grid bias. Anode current then flows, causing the relay to operate and change the speed of the motor. The size of the loop is now reduced, but the process is checked when the tape is pulled clear of the contact, for the bias is thus reapplied to the thyratrons, causing the relay to release and restore the original condition.

The constant-speed drive system is not required during the rewind, and an automatic release clutch is provided for disconnecting it. The rewinding motor, running in the reverse direction and coupled to the tape-drive through a double-speed pulley, pulls the tape off the right-hand reel and feeds it into the large tape reservoir. The winding motor, with its direction of rotation reversed, now drives the left-hand reel on to which the tape is being rewound, and its speed is regulated as before by the thyratron circuit. The speed of the unwinding motor is not varied during the rewinding operation.

The trailing reel in both directions of working is restrained by a simple band-brake, and provision is made for increasing the brake tension in order to overcome the momentum of the reel when the machine is stopped.

The operation of the machine is controlled by a rotary gate switch which has five positions:—Off, 1st and 2nd wind, and 1st and 2nd rewind. The intermediate positions are introduced to ensure the correct operating sequence, and to prevent the formation of loops of tape in the reservoirs during the processes of starting and stopping.

The behaviour of the machine for the various positions of the gate switch is given in the schedule on page 90.

The performance of this machine shows a considerable improvement on that of earlier designs. The variation in tape velocity is not normally perceptible when listening to a musical recording.

Illustrations and references showing the general construction of the machine are given in Figs. 19–22 (see Plates 1, 2, 3, and 4).

SECTION (4)

Magnetic Recording Applied to Broadcasting

As the technique of broadcasting has developed during the last few years, so the demand for recording

SCHEDULE OF OPERATIONS

Position of Gate Switch.	**Action.**
Off.	Power is applied to the heaters of the thyratron cathodes, and to the metal-oxide rectifier circuit providing the grid bias.
No. 1 *Wind.*	H.T. is switched on to the thyratrons and the L.T. and grid-bias supplies continued. No. 1 motor is switched on for forward running at its lower speed, engages with the belt drive of the R.H. reel by means of No. 1 automatic dog clutch, and tends to drive this reel in an anti-clockwise direction. Maximum braking is applied to the trailing (L.H.) reel.
No. 2 *Wind.*	The conditions of the *No.* 1 *Wind* position are maintained, except that the braking on the L.H. reel is reduced to the normal running value. No. 2 motor runs forward at its higher speed, engages with the lower-geared belt system through No. 2 automatic dog clutch, and thus causes No. 2 tape drive to pull the tape off the L.H. reel and feed it forward at a rate slightly higher than the feed of the main tape-drive system. No. 3 motor is switched on and runs forward at its constant speed of 250 r.p.m., pulling the tape through the heads at a steady rate of 90 metres per minute and feeding it into No. 1 reservoir from the R.H. side. The higher speed of No. 2 motor is maintained until the tape loop in No. 2 reservoir touches the thyratron contact, thereby short-circuiting the grid bias and causing the operation of No. 2 relay. The speed of No. 2 motor is thus changed to the lower value, which allows the tape loop in reservoir No. 2 to be reduced. The lower speed is maintained until the loop ceases to make contact with the thyratron control contact, whereupon the motor reverts to its higher speed and the process is repeated. The R.H. reel, driven by No. 1 motor and now having no tape tension to restrain it, begins to wind up the tape, but at a lower rate than it is being fed into No. 1 reservoir by No. 3 system. However, when the loop in No. 1 reservoir touches the thyratron contact, No. 1 motor runs at its higher speed and continues to do so until the tape loop in No. 1 reservoir is pulled off the contact. The speed then returns to the lower value and the process is repeated.
No. 1 *Rewind.*	H.T. is switched on to the thyratrons and the L.T. and grid-bias supplies are continued. No. 3 drive system clutch is disengaged. No. 1 motor is switched on for reverse running at its lower speed, engages with the higher-geared belt drive of the L.H. reel by means of No. 1 automatic dog clutch, and tends to drive this reel in an anti-clockwise direction. Maximum braking is applied to the trailing (R.H.) reel.
No. 2 *Rewind.*	The conditions of the *No.* 1 *Rewind* position are maintained, except that the braking on the R.H. reel is reduced to the normal running value. No. 2 motor runs backward at its higher speed, engages the higher-geared belt system through No. 2 automatic dog clutch, and thus causes No. 2 tape drive to pull the tape off the R.H. reel and feed it into No. 1 reservoir from the L.H. side at approximately twice the normal forward speed. No tape loops occur in No. 2 reservoir, hence the speed of No. 2 motor is not controlled. The L.H. reel, driven by No. 1 motor and relieved of restraint by the existence of the loop in No. 1 reservoir, then begins to wind up the tape at a low rate until the loop in No. 1 reservoir touches the thyratron contact. When this contact is made No. 1 motor runs at its higher speed and continues to do so until the loop in No. 1 reservoir is pulled off the contact. The speed then returns to the lower value and the process is repeated. No. 3 motor does not run at all during the rewinding process.

facilities has increased. New ideas of presentation have materialized and the inauguration of the Empire service called for some means of providing programmes, other than that of direct transmission from a studio.

Three different systems of recording are in use at the present time, as it is impossible to fulfil all the technical and programme requirements with any single system.

The programme service requirements are many and varied. It must be possible to make a record of a programme lasting for some hours, take extracts from it, mix, fade, and superimpose extracts from other recordings, and produce what is termed a composite record. It is also necessary, in some instances, to make a number of copies of a record.

The types of recordings can be divided into two groups—permanent and non-permanent.

Permanent recordings are made of items of historic interest, such as national ceremonies and speeches; effects records used to provide backgrounds in the production of plays, and so on.

The non-permanent recordings include those of rehearsals, and of home programmes for subsequent transmission on the short-wave Empire service. This may involve four or five reproductions to cover the various zones, after which the record is of no further interest and may be destroyed.

The technical considerations include:—

(1) All factors involved in the production of what is often termed " good quality."

(2) The possibility of the duplication of the original record without noticeable deterioration of quality and without resorting to re-recording.

(3) The time which must elapse before the first reproduction can be obtained.

(4) The reliability of the system as a whole.

(5) The reliance which can be placed on the success of any individual recording.

(6) The ease of operation of the system.

(7) The permanence of quality of a record, (*a*) during storage and (*b*) after repeated reproduction.

(8) The suitability of the records for storing from the point of view of weight, size, etc.

If consideration is given to the matter, it will be seen that magnetic recording on steel tape fulfils a number of the programme and technical requirements. It is very useful in cases when non-permanent records are required, as the tape can be used repeatedly. It is also applicable in the case of permanent recordings, as the quality does not appreciably deteriorate after repeated reproductions. The records are, however, large and weighty, and occupy a considerable space when stored.

The system is particularly suitable for recording programmes of long duration which are required for subsequent reproduction in their entirety. It is not convenient, however, in cases where it is known that a number of extracts will be required from the record to form a shortened version of the original, since it is undesirable to cut and join the tape. This difficulty can be overcome to a certain extent by re-recording, but a slight deterioration of quality is likely to occur as a result of this process. Furthermore, it is generally more convenient to employ a recording system which is particularly applicable to this type of work.

Description of a Tape Recording Channel

Each recording channel at Maida Vale comprises two machines and two complete sets of amplifiers and associated apparatus, which are accommodated in two rooms. One is known as the apparatus room and contains two machines, amplifier bays, associated apparatus, and a loudspeaker for monitoring purposes. The other is known as the control cubicle, in which there is a second loudspeaker and a control desk from which the control engineer can see into the apparatus room through a double glass window. A schematic diagram is shown in Fig. 23, to which the references apply.

The programme originating in a studio or at an outside broadcast point is transmitted by line via the control room at Broadcasting House to the control room at Maida Vale. It is then distributed by isolating amplifiers to the recording channels. The programme is next passed through a control potentiometer in the control cubicle, and thence to the two recording amplifiers (Rc) the inputs of which are connected in parallel through a key switch (Rec. C.O.), which can be used for change-over purposes. Each amplifier incorporates suitable frequency-discriminating networks between the first two stages, and has a final push-pull stage. The output of the amplifier can be connected by means of a switching key to either of the recording heads associated with the appropriate machine, the steady demagnetizing current being introduced into this circuit.

Either of the two reproducing heads associated with the machine can be connected to the input of the three-stage reproducing amplifier (Rp) by means of a key. This amplifier embodies the discriminating networks necessary to produce an overall frequency characteristic that is sensibly flat up to the cut-off frequency.

Two separate output circuits are provided. One of these is used to supply the programme to the Maida Vale control room through a change-over unit (2 Ch. Fade Unit) and thence to Broadcasting House. The other supplies the loudspeaker amplifiers associated with the loudspeakers (L.S.) in the apparatus room and the control cubicle. Switching arrangements are available in both these rooms, to enable the loudspeaker to be connected either across the incoming programme or across the output of either of the two reproducing amplifiers. The listening levels can be balanced by means of volume controls. A programme-level indicator (P.M.) is available for measurement of the level at any point in the chain.

Since the recorded programme is reproduced a fraction of a second after the recording actually takes place, the quality of this can be readily compared with that of the incoming programme. It has already been pointed out that this facility enables the recording engineer to determine and adjust the recording level by listening to the programme after reproduction, since it is necessary to effect a compromise between limits imposed by signal/noise ratio and distortion.

The above method of monitoring has been found to be the most satisfactory in practice, but various forms of visual indicators have been tried from time to time. One of these consisted of a neon indicator operated indirectly by the a.c. voltage developed across a resistance in series with the recording head. This indicator was so adjusted that faint flashes occurred in the lamp when the maximum permissible recording level was approached, and bright flashes when the level producing overloading was reached. There are, however, certain drawbacks to such a system. It must be realized that no reduction of the value of available signal/noise ratio can be tolerated. In fact, in some instances it is necessary when recording to reduce the volume range of the programme which has already been controlled for transmission purposes.

In order that an indicator of the type described above should be of assistance, the tapes used for recording must be absolutely uniform as regards the maximum permissible recording level which can be applied. If this is not the case, the visual indicator must be so adjusted that a margin is allowed for the non-uniformity of the tapes, and this means that the maximum permissible recording level is not supplied to some tapes, with a consequent reduction in signal/noise ratio.

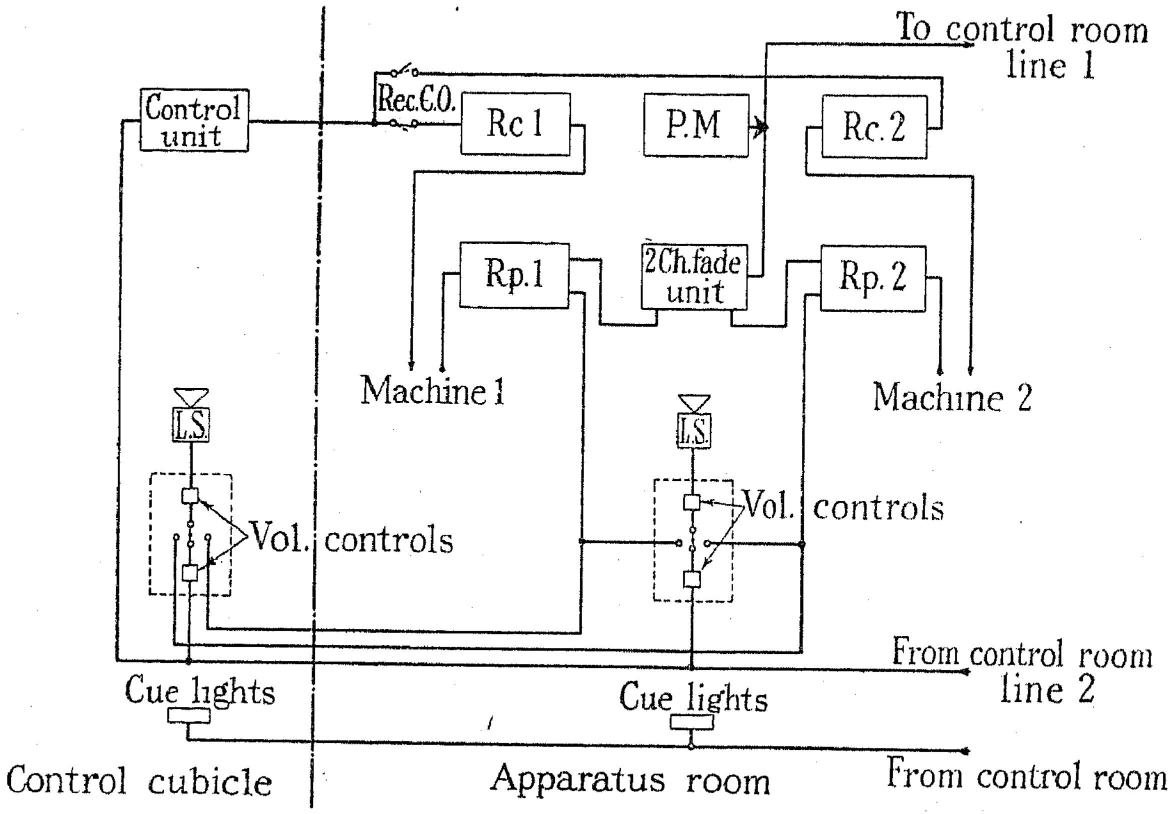

Fig. 23.—Block schematic of recording channel.

Two engineers are required for each recording channel, one in the control cubicle and the other in the apparatus-room. The former checks the quality of the recording and controls the recording level, while the latter is in charge of the machines and apparatus. A system of illuminated indicators is used between the control cubicle and the apparatus room, which enables the control engineer to give the necessary directions to the engineer in charge of the apparatus.

When preparing for a recording session, the pole-pieces for all the heads must be carefully selected, and a test recording must be made to ensure that the performance is satisfactory.

Both machines in the recording room are kept ready for immediate use, whether the recording to be carried out requires one spool of 30 minutes' duration or more. In the case of a recording lasting over 30 minutes it is necessary to make a change-over from one machine to the other. This is usually carried out in the following way. When the spool on the first machine nears the end of its run the second machine is started up, and at a suitable passage in the programme, preferably during a pause between two items, the recording is changed from the first machine to the second. The tape on the first machine is then rewound, and a new tape put into position ready to take over if necessary. In the event of a fault of any kind developing while the first machine is in operation, the second machine is immediately started up and takes over the recording. If, on the other hand, a fault develops on the second machine while the first is being rewound, it is necessary to continue the recording on a machine from another channel. No change-over is attempted in this case, the spare machine being brought into operation as soon as possible.

In the case of a reproduction involving more than one spool of tape, the machine allotted for the reproduction of the second spool is set up in such a way that there is just enough tape prior to the commencement of the recorded programme to allow the machine to reach normal speed before the recorded section reaches the reproducing head. An appropriate passage is selected towards the termination of the recording on the first spool, to act as a cue for the starting of the second machine. The finding of cue passages is simplified by the provision of indicators coupled to No. 2 driving system (Figs. 19 and 21, Ref. 49, Plates 1 and 3).

When a programme is of particular importance, simultaneous recordings are made on two machines. These may be the two machines comprising one channel when the programme is timed to last for less than half an hour, otherwise it is necessary to use two complete channels.

The signal for the commencement of a recording or reproduction can be given by a system of cue lights, which are operated either from the control rooms at Maida Vale or Broadcasting House, or directly from a studio. Alternatively, the recording or reproduction can be arranged to commence on a predetermined cue-passage occurring during a normal transmission.

It is sometimes necessary to obliterate a certain portion of a recording. For instance, in the case of a programme recorded during the day for transmission to the Empire at night, an announcement such as " This is the London Regional Programme " may occur, and before the recording is transmitted to the Empire this must be wiped out. This is accomplished by inserting a wiping pole-piece in the second reproducing head and using it as a wiping head. A unit consisting of a battery and a variable resistance with a spring-loaded control is connected in series with this head. The resistance is open-circuited in the " off " position of the control, and a condenser is connected across it to prevent clicks due to sudden changes of flux. The tape on which the passage to be wiped occurs is passed through the machine with the first reproducing head in use in the normal manner. At the appropriate moment, the control completing the circuit through the second reproducing head is operated and the unwanted portion is wiped out. At the end of this unwanted portion the control is released and the remainder of the recording is unaffected. It is then possible to record an alternative announcement on the portion of the tape which has been wiped. By this means the wiping process may be commenced practically instantaneously or the programme may be faded out gradually.

BIBLIOGRAPHY

E. HORMANN: " Concerning the Theory of Magnetic Sound Recording," *Elektrische Nachrichten-Technik*, 1932, vol. 9, p. 388.

E. MEYER and E. SCHULLER: " Magnetic Sound Recording on Steel Tape," *Zeitschrift für technische Physik*, 1932, vol. 13, p. 593.

H. J. VON BRAUNMUHL: " Sound Recording in Broadcasting," *Funktechnischer Vorwarts*, 1935, vol. 5, no. 4.

C. N. HICKMAN: " Sound Recording on Magnetic Tape," *Bell System Technical Journal*, 1937, vol. 16, p. 165.

ACKNOWLEDGMENTS

In conclusion, the authors wish to thank the Chief Engineer of the British Broadcasting Corporation and the Research Manager of Marconi's Wireless Telegraph Co. for permission to publish this paper, which is largely based on research and development work undertaken by the Corporation and Marconi's Wireless Telegraph Co.

Thanks are also due to Mr. D. B. Weigall, B.A., of the British Broadcasting Corporation's Research Department, for valuable assistance in the preparation of the paper, to Mr. N. M. Rust of Marconi's Wireless Telegraph Co., who was in charge of the preliminary development work, and to Messrs. James McMillan and Co., representing the Uddeholms Co. of Sweden, for information concerning the manufacture of the tape.

DISCUSSION BEFORE THE WIRELESS SECTION, 1ST DECEMBER, 1937

Mr. N. M. Rust: While an ideal recording system does not exist, there are several features of magnetic recording which stand out. There is no other recording system, for instance, whereby it is possible to " rub out " in the middle of a programme and to record something else. With magnetic recording one also has the ability to play back instantaneously and to make a delay playback at any time from 0·1 sec. up to 1 minute after the signal.

The ability to rub out and re-record could be applied to recording for stenographic purposes. There is the possibility in a large organization of a recording room linked up with the offices and the typists in such a way that the staff can record their letters, have them played back again, and then rub out and correct them. Whether such a system is commercially practicable at present I do not know, but if it could be developed it would certainly be a great boon in large offices.

With regard to the ability to play back at any delayed time, one fairly obvious application which seems quite a simple matter is the production of artificial echoes for broadcasting. But, if it were done in the ordinary way with a single reproducing head, the system would not give a natural effect, because such an arrangement does not simulate the reverberations which occur. The only feasible system would be one having two or three reproducing heads delayed at different times behind the recording head, and with the delay periods arranged in relation to the reproduction characteristics. The heads and the amplifiers associated with them would have to be split into bands, the head nearest the recorder passing the higher frequencies and the head farthest away passing the lower.

In developing the recording machine to the point where speed variations were not noticeable in commercial production, even higher standards were obtained under laboratory conditions. One of the difficulties of magnetic recording is that mechanical considerations necessitate a certain size of tape drive wheel. If the wheel is made smaller it does not hold the tape well, and if it is made bigger the inertia effects in the tape wheel itself become serious. It happens that the period of variation which the ear most notices is just about the period of one revolution of the tape wheel, which makes it necessary to take great precautions to ensure a steady drive.

By using a synchronous direct-drive motor instead of a high-speed motor and gearbox, we avoided " flutter " troubles. The electrical analogue of such an arrangement is a generator working through an inductance. The flexible coupling interposed between the direct drive and the flywheel is equivalent to a condenser, and the flywheel itself is equivalent to a big inductance. The inertia of the tape-wheel system is equivalent to a small inductance, and the aim is to keep this small relative to that represented by the flywheel, and also to make the flywheel equivalent to a damped inductance. Similarly, by using material of high internal friction, we attempted to get all the damping we could in the coupling.

In the experimental work to get rid of the troubles due to irregularities, we used two forms of check arrangements. The first consisted of a stroboscopic disc on the end of the tape wheel, and a sodium lamp which was energized from the mains and timed from a special circuit. The stroboscopic marks stood out so clearly that it was possible to use a low-power microscope for observing the fluctuations. The other method used for the final check was an oscillograph method, suggested by the B.B.C. engineering staff. It consisted in controlling the horizontal sweep of the oscillograph by a sawtooth time-base which was locked from the recorded tone, and obtaining the vertical sweep from the reproduced tone. Any irregularity in the speed caused a displacement of the pattern.

An interesting piece of apparatus used in the development work was a continuous-band tape machine. This had a form of reservoir which allowed of the use of a tape taking up to ½ minute to run through. The tape had a joint in it, of course, and was so arranged that we could easily reverse it. This device enabled us to listen to speech backwards. We would take a simple sentence and try to say it backwards; if we were successful, the record would come out as the original sentence, which seems to show that intelligibility of speech, at least in English, is not dependent on transients (which surely cannot be generated backwards) to the extent that one would think.

Mr. H. L. Kirke: Some of the effects of demagnetization are not easy to understand, and I should welcome further explanation of these.

I should like to mention the difficulty of getting steel manufacturers interested in the problem of steel tape for magnetic recording. The trouble is that the cost of research in this subject is very high, and relatively only a small quantity of tape is required. For this reason we have more or less had to adhere to the products of one firm. Another point on which the authors do not say very much is the question of the pole-pieces. Further details of these would be welcome.

Regarding the subject of jointing, it would be interesting to know how the welding experiments referred to in Section (3) have turned out.

In connection with constancy of speed and size of apparatus, there is a great difference between the requirements of an apparatus for reproducing music in ordinary broadcasting channels and that of apparatus which is merely required to reproduce speech. Speed constancy in music requires far greater accuracy than in speech, and it is necessary to use much larger apparatus in order to obtain that speed constancy. For speech records in offices and the like, a much smaller machine than the one described in the paper would be satisfactory.

Mr. F. H. Dart: I consider the machine described in the paper to be a reliable piece of mechanical apparatus. In the past year machines of this type have been " on the air " for reproducing programme material 1 957 hours. The total breakdowns over the same period only amount to 39 minutes 28 seconds, equivalent to 0·033 % of the programme time. Of these breakdowns, none can be directly attributed to the machines: 26 minutes 33 seconds of the breakdown time was due to tapes breaking at joints, the remainder being due to amplifier or other electrical trouble.

Regarding the authors' statement that the single-pole and the double-pole methods of working can both be made to give satisfactory results, I would recommend the use of the single-pole method. In the past our tape joints have taken the form of soft-solder lap joints, and it is my opinion that breakdown has resulted from the pole-pieces catching in these joints. With the single-pole system, the edge exposed to the pole-piece can be so arranged that it is on the opposite side of the tape from the pole-piece, and in this way the trouble can be materially reduced.

No mention has been made of the wear of tapes. We have many in service which have carried out 60 recordings; they must have been reproduced many hundreds of times, and yet they are still in good condition. A tape which was delivered from the manufacturers just over a year ago, and on receipt was calibrated both for background noise and magnetic properties, after a year's service and 30 recordings showed an increase of 6 db. in background noise. This was first thought to be due to general wear on the tape, but was subsequently found to be due to foreign matter, and by suitable cleaning the tape was restored to its original state.

There is one question I should like to ask the authors. The recording curve A in Fig. 16 (double-pole system) is relatively flat compared with the recording curve A in Fig. 17 (single-pole system), whereas the reproducing curve C in Fig. 16 (double-pole system) has a sharply-rising characteristic while the corresponding curve for the single-pole system (Fig. 17) is a falling one. It would therefore appear that the background noise on a double-pole system should be materially greater than on a single-pole system, and I should be interested to know whether this is the case.

Dr. L. E. C. Hughes: While working on the development of a magnetic reproducing machine, which consisted of a steel tape clamped on the edge of a wheel, my associates and I obtained some 12 000 reproductions before the level dropped by about 3 db.; the intelligibility was maintained, but the quality was only of telephonic standard. In this machine, an arrangement provided for the lifting of the pole-tips from contact with the tape on the passage of a joint.

With ordinary steel-tape we noticed a threshold effect, i.e. a definite applied level had to be exceeded before any registration was effected; also with increasing applied level the registration level fell off, on account of saturation. Did the authors experience such effects in the early stages of development?

I wish to protest against the physical explanation of demagnetization of steel, using the theoretical conception of magnetic poles. These do not exist physically, and are useful only for defining an arbitrary magnitude of magnetic flux, which, like its electric counterpart, current, is a sufficient conception for explaining magnetic phenomena.

Mr. Donald McMillan: Five years ago we in the Post Office Engineering Research Department acquired a magnetic-recording apparatus which we employed to introduce a time-delay into transmission circuits. It had certain differences from the equipment described in this paper, and I should be interested to have the opinion of the authors on those differences. The arrangement of the equipment was simple. It consisted of a continuous loop of tape which passed first through a recording head, then through a reproducing head, and finally through a wiping head; and the time-delay could be controlled by moving the recording head relative to the reproducing head.

The method adopted for wiping out was, however, quite different from that described by the authors. The wiping head consisted of a solenoid with its axis along the tape, and the tape was entirely demagnetized by passing alternating current through this solenoid. The tape was then magnetized to a certain steady value by the recording head, which had direct current passing through it. The recording was presumably effected by alternating the value of the magnetization around the value produced by that direct current. In the system described by the authors, on the other hand, the wiping head saturates the tape, and the recording in effect reduces the magnetization of the tape.

I am interested in the description of the present apparatus from the point of view of the signal/noise ratio and range of amplitude permissible. A recently-published description of a good type of optical recording system states that a ratio of peak-to-valley intensity in music of 55 db. was reduced effectively to 38 db. by the use of this recorder. Can the authors quote any comparable figure for their equipment?

With regard to the signal/noise ratio, a figure of 35–45 db. is considered poor by telephone engineers, whereas the reproduction which we have heard from the author's apparatus and with which we are familiar from broadcasting is anything but poor. I should like to know the explanation for this. Further, I am not able to relate the figure of 35–45 db. to Fig. 14. Perhaps the authors could give some further information as to how that signal/noise ratio was measured. I presume that it was measured with some form of programme meter, measuring first of all the signal and then the noise.

With our old type of apparatus, which used rather thicker pole-pieces than those described in the paper, we had great trouble due to wear of the pole-pieces. I notice that the pole-pieces described by the authors are sharpened down to 0·06 mm., and I should like to know what is the effect of wear upon their performance. Presumably the degradation obtained as they wear is the conventional type associated with any recording system, namely lack of high frequencies and increase in noise. Does this constitute a serious cause of trouble, and do the pole-pieces need changing often?

Mr. L. S. Anand: Can the authors give any information about the design and the thickness of the tape? I imagine that while there is longitudinal recording there is also transverse recording, and this may distort the regularity of reproduction. Is the thickness of the tape based on mechanical limits, and should the tape be as thin as possible? Is there an optimum value for the coefficient of modulation used?

Mr. L. C. Stenning: In view of the advantage derived in the way of cost and storage by reducing the speed of the tape, I should like to draw attention to some recently-published American work* in which the tape was magnetized perpendicularly instead of longitudinally. The quality and frequency-range claimed were about the same as for the authors' machine, and this result was secured with a speed of about 16 in. per sec., instead of 1·5 m. per sec.

Mr. E. L. E. Pawley: We all know the advantages of negative feed-back in an ordinary amplifier system: one takes part of the output and feeds it back into the input, and, if the phases and magnitudes are right, one is thus able to cancel out some of the unwanted effects which occur in the process of amplification. In the case of magnetic recording, would it be possible to take part of the output from a reproducing head and inject it back into the recording head (or into an auxiliary recording head) during the recording process in such a way as to cancel some of the non-linear distortion and also some of the noise which arises in recording? An obvious difficulty is that the phases will not be the same for all frequencies, but it might be possible to put this right by a phase-compensating network or by running an auxiliary loop of tape in synchronism with the main tape.

Mr. H. Bishop: I should like to mention the use of magnetic-recording equipment for business purposes. The original machine of Dr. Stille was developed by him and marketed as an office machine in 1925, because at that time the need for a recording process for broadcasting had not asserted itself. The machine had two forms. One used a steel tape 6 mm. wide, compared with 3 mm. on the machine described in the paper, and the other, a much smaller affair, used a wire of 0·25 mm. diameter. Both gave what would now be called very poor quality, but in those days the results were considered not at all bad. No further development work was done on the machine until it became available in this country in 1930.

More recording in connection with broadcasting is done in Germany than in this country. The German broadcasting authorities use magnetic-recording machines installed in vans, and the results are, I believe, fairly successful. In this country the vans used by the B.B.C. are equipped for cellulose-disc recording.

An important requirement of a recording system when used for broadcasting is the ability to edit the records. For example, during a football match a commentary of an hour or so may be recorded, but only the specially exciting parts may be wanted for re-broadcasting in the news bulletin the same evening. For this class of work, therefore, editing is essential. A magnetic-recording system is not quite so convenient in this respect as a disc system, but on the other hand the former has the advantage of being able to repeat a long programme without interruption. Further, the tape can be used an indefinite number of times, and this is of particular value for the short-wave service, where the programmes are radiated 4 or 5 times in a period of 24 hours, after which they are not wanted again.

Messrs. A. E. Barrett and **C. J. F. Tweed** (*in reply*): The experiments relating to the welding of tapes to which Mr. Kirke refers have proved satisfactory. During the process of welding, however, the tape becomes brittle and subsequent annealing is necessary. This unfortunately alters the magnetic properties of the tape. The noise due to this and to the mechanical imperfections of the joint is considerably less serious than that produced by a soldered joint, which is comparatively uneven and mechanically inferior.

Mr. Dart's remarks concerning single-pole recording (Group 1, Fig. 11) and double-pole recording (Group 2*b*, Fig. 11) are important since they have considerable bearing on the operation of a system with which he is concerned. There is little difference between the results obtained with these two systems, but listening-tests indicate that the quality of reproduction is superior when using the double-pole system. With the latter it is possible to record at a slightly higher level at the higher frequencies, though at the lower frequencies the converse holds. As we are chiefly interested in the level of the recording at the higher frequencies from the point of view of surface noise, this point is in favour of the double-pole system.

---

* *Bell System Technical Journal*, 1937, vol. 16, p. 165.

The statements referring to the recording curves for the S.P.I. and D.P. systems appear correct at first sight, but it must be remembered that the signal/noise ratio is dependent on the absolute level of recording, and in practice it is found that the increase in signal/noise ratio which might be expected is not realized. The results of tests under service conditions show that for a given quality of reproduction the recording level for the S.P.I. system is lower than the corresponding level for the D.P. system. This results in practically the same signal/noise ratio being obtained for both systems.

It is probable that, during the experiments referred to by Dr. Hughes, recording was carried out near the state of zero magnetization, in which case the effect mentioned might be expected on account of the low initial permeability of the tape. We regret that Dr. Hughes takes exception to our physical representation of the demagnetization of steel.

The method of wiping mentioned by Mr. MacMillan indicates that the tape is in a state of zero magnetization when leaving the wiping head, and it follows, therefore, that the recording takes place on a portion of the initial magnetization curve instead of a portion of the largest hysteresis loop, and ideally a shorter straight portion would be available for the recording process. Since there are other limitations to the possible amplitude of recording, the linear ranges obtainable with the two systems may not be widely different. With regard to a comparable figure for peak-to-valley intensity which may be obtained, this varies considerably with different tapes, but is of the order of 30 to 35 db. The signal/noise ratio was measured with various types of programme meter. Some of these meters had a long, and others a short, charging time, but the results obtained agreed within one or two decibels. Measurements on the tapes used at the demonstration actually gave a signal/noise ratio of approximately 40 db. The energy contributing to noise is chiefly confined to the high frequencies, and if measurements were made with a cut-off frequency of 3 000 cycles the signal/noise ratio would be higher. If an attempt is made to estimate the maximum signal/noise ratio from the curves of Fig. 14, allowance must be made for the amplitude/frequency relationship which occurs in music, and since the amplitudes at 4 000 cycles are considerably lower than those occurring below 1 000 cycles it would appear reasonable to estimate the signal/noise ratio from one of the lower-frequency linearity curves. The curves in Fig. 14 were not taken with a particularly good tape—probably one which gave a signal/noise ratio of 36 db. The effect of pole-piece wear is in general to reduce the high-frequency response, and also the noise, since this is noticeable chiefly at the higher frequencies. If, however, the pole-pieces become jagged due to passing over a bad joint, then the quality is liable to suffer. Pole-pieces are checked before each recording or reproduction, and their life is largely dependent upon the nature of the joints in the tape.

In reply to Mr. Anand's questions concerning the thickness of the tape, this dimension (0·08 mm.) was originally determined by considerations of mechanical strength and manufacture. Experiments have been carried out recently using tape of a thickness of 0·05 mm., but the results have been inferior to those obtained with the 0·08 mm. tape. The expected reduction in self-demagnetization effects was not realized, owing probably to non-uniformity of magnetization at the higher frequencies referred to in Fig. 15.

We have carried out experiments on a system similar to that referred to by Mr. Stenning, and the results were inferior to those obtained with the D.P. or S.P.I. systems. We found that the adjustments which were necessary to achieve satisfactory results were much more critical, which is an added complication from the point of view of a recording service.

The suggestions outlined in Mr. Pawley's remarks are very interesting, but their application presents serious practical difficulties.

A THEORY OF FLUCTUATION NOISE*

By D. A. BELL, B.A., B.Sc., Graduate.

(*Paper first received* 28*th June, and in final form* 26*th November*, 1937; *read before the* WIRELESS SECTION 5*th January*, 1938.)

SUMMARY

From a consideration of the physical phenomena involved in thermionic conduction and in thermal or Johnson noise, it is deduced that the noise in a space-charge-limited valve is best expressed as a thermal noise, and it is shown that a small correction must be applied to the valve slope resistance to give the value of resistance effective as a noise source. The theoretical temperature of this resistance is then shown to be approximately half the cathode temperature. In a temperature-limited valve, both the resistance and the temperature of the conducting path are indeterminate. The "pure shot noise" formulae are applicable to this case, and represent the maximum noise which can result from the passage of a given current through the system. It is incorrect, however, to state that in space-charge-limited conditions "pure shot noise" is smoothed out, while thermal noise in the internal resistance of the valve appears as an additional factor; shot noise and thermal noise in the valve's internal resistance are essentially the same phenomenon, but are modified by the differing conditions of electron transit.

Nyquist's expression for thermal-agitation noise is derived from the atomic mechanism in the case of a metallic conductor.

(1) INTRODUCTION

The author might perhaps be accused of temerity in attempting a theory of such weighty subjects as shot noise and thermal-agitation noise without having the backing of any experimental work; but from discussion with workers in this field there seemed to be scope for one with training as a physicist to endeavour to interpret the results of other radio workers, even though statisticians regard the problem as unpromising. Those who are familiar with the subject are asked to forgive the elementary treatment of simple shot noise with which the paper opens; it is included because it appears to be an appropriate starting point in the search for a coherent theory based on fundamental principles.

Although in the realm of mechanical engineering we do not find the atomic structure of matter forcibly brought to our notice, yet in electrical technique we have already reached the stage where the discrete nature of the electric current is clearly perceptible. As an extreme instance of the sensitivity attainable by thermionic-valve methods, an electrometer valve can be made to detect a current of the order of 10^{-18} ampere—say 10 electrons per second. If the indicating system could be made of sufficiently short time-constant, such a current would necessarily be found to be non-uniform, but other information is required to predict whether the variation will be a regular 0·1-second cycle, or an irregular distribution subject only to the constancy of the long-period mean. (Actually, these minute currents can at present be measured only by a condenser-charging method, so we are very far from the required rapid indicating system which would give the direct experimental answer to this question.) Considering only electron currents (i.e. disregarding any phenomena such as electrolysis or gas discharge, which depend in part on conduction by heavier ions) we may distinguish between electrons travelling within a conductor and those travelling through free space; in the problem of valve noise we are concerned with the latter category. Now electrons may be emitted from conductors into space by various mechanisms—radio-activity, thermal emission, photo-electric emission, and ionic or electronic bombardment (the last including secondary emission). Atomic physics shows that all these processes are essentially "random," meaning that although it is possible to predict the mean value of any quantity involved, from statistical laws based on previous observations of large numbers of events—just as actuaries will determine a statistical value for the length of life of any class of persons—yet the behaviour of the individual unit, when we come down to atomic phenomena, is as incalculable as the life of an individual human being.

It is reasonable to assume, therefore, that the electrons constituting the currents in a thermionic valve are emitted at random times. Suppose, then, that five different currents, each having a mean value of 10 electrons per second, could be observed for 2 seconds, and that the 20 electrons in each were found to be distributed over the 2 seconds in the manner represented by the curves lettered A, B, C, D, E, in Fig. 1, where the 2-second period has been divided into intervals of 1/10th second. In each current the mean number of electrons is 1 per 0·1-second interval, but the individual values vary between 0 and 4 electrons per interval; while adding all five currents together, giving a mean value of 5 electrons per 0·1 second, is seen from the curve marked "Sum" to produce variations from 2 to 12. Although this diagram has not, of course, any great quantitative significance,† it clearly illustrates that as the number of electrons with which we are dealing becomes greater (i.e. the current is increased) the rate of flow becomes *relatively* smoother, but the *absolute value* of the variations is larger. Thus, on the one hand, large currents are for ordinary purposes regarded as

---

[*] Reprinted from *Journal I.E.E.*, 1938, vol. 82, p. 522.

[†] The random distributions of Fig. 1 were obtained by drawing lots as follows. Cards numbered 1 to 20, corresponding to the 20 intervals of the period plotted, were shuffled and draws were then taken at random; after noting the number each card was replaced and shuffled, the process being repeated until 100 numbers had been drawn. The first 20 of these numbers, in order of drawing, were taken to represent the numbers of those intervals of the period during which an electron passed for current A, and successive batches of 20 for the other currents. The number of electrons attributed to the *n*th interval of 0·1 second is thus equal to the number of occasions on which a card numbered *n* was drawn during the 20 draws representing the particular current. The result should be genuinely "random."

uniform, while on the other hand the shot noise in a temperature-limited valve increases in proportion to the current. If in a wireless receiver the steady current of the valve in each stage of amplification were in proportion to the signal strength, so that the anode current in each valve was approximately 100 % modulated by a strong signal, this type of phenomenon would still cause a deterioration of signal/noise ratio as the sensitivity was raised, since the smaller currents in the input circuit would then be subject to a large relative fluctuation; this corresponds to the case of a photocell combined with an electron-multiplier. The increase of noise here arises from the fact that with decreasing signal amplitude the charge conveyed across a valve per half-cycle of signal current is becoming more comparable with the charge on a single electron. But in a normal amplifier one might expect a far worse state of affairs, since the current in the first valve depends upon the type chosen, and the size of valve cannot be indefinitely reduced in proportion to the signal strength. Actually, however, the ordinary thermionic amplifying valve works under conditions of space-charge limitation (not temperature limitation), and it is known that the valve noise is then less than predicted by the simple shot-noise theories. The noise in the presence of space-charge limitation is discussed later in this paper.

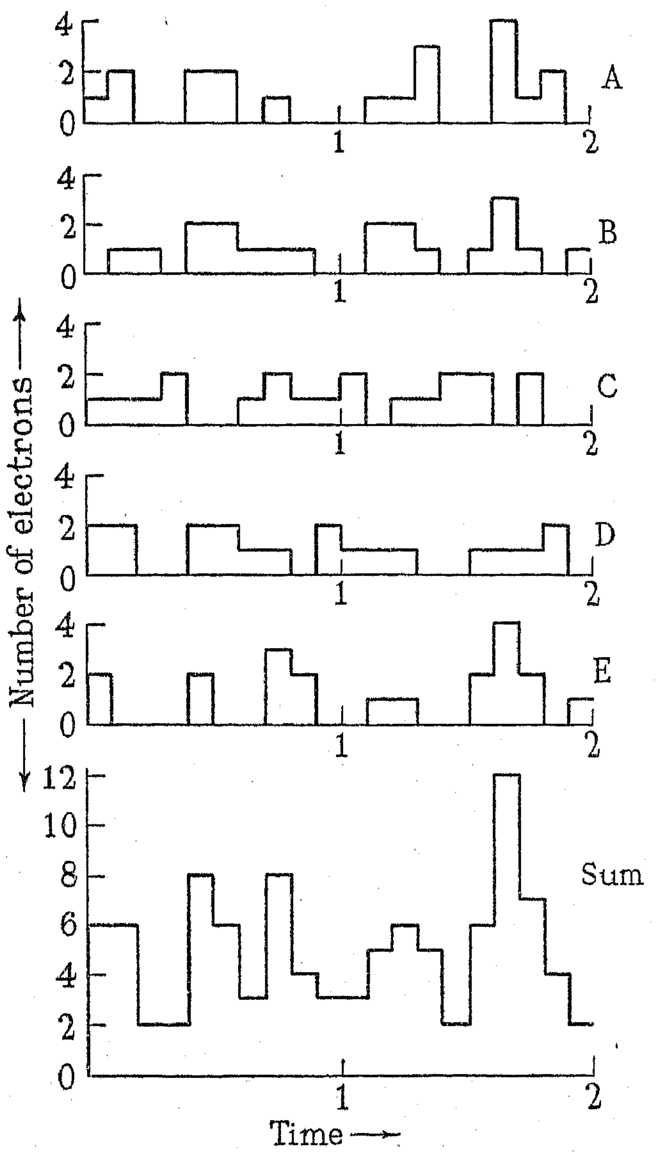

Fig. 1

(2) TEMPERATURE-LIMITED CONDITIONS

Existing theories, which are well supported by experimental results in the temperature-limited regime, show that the mean square of fluctuation voltage at the anode of a temperature-limited diode is directly proportional to the magnitude of the current from which it arises. Moullin and Ellis,* for example, give the formula

$$V_s^2 = I_{av.}eR/(2C) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where V_s^2 is the mean-square shot voltage in an anode circuit, having resistance R shunted by capacitance C, connected to a diode having temperature-limited anode current of mean value $I_{av.}$. For an oscillatory anode circuit they give

$$V_s^2 = \frac{I_{av.}eL}{2RC^2}(1 + F^2) \quad . \quad . \quad . \quad . \quad (2)$$

where $F = R/(\omega_0 L)$. Now for a good tuned circuit $R/(\omega_0 L)$ at resonance ($\omega_0 = 2\pi$ times the resonant frequency) is of the order of 10^{-2}, so that F^2 is negligible compared with unity. Writing $L/(RC)$, the dynamic resistance of the circuit, as R', the formula for noise in a tuned circuit becomes

$$V_s^2 = I_{av.}eR'/(2C) \quad . \quad . \quad . \quad . \quad (3)$$

which is of the same form as (1), but employing the dynamic resistance of the circuit in place of an ohmic resistance. It is interesting to note that in (3) the factor C still appears in the denominator. This suggests that the use of a tuned circuit of low L/C ratio with low resistance will give the best signal/noise ratio for a given amplification, i.e. for constant value of the dynamic resistance R'. However, this policy is limited firstly by the difficulty of lowering the resistance sufficiently to maintain the desired value of dynamic resistance, and secondly by the increasing selectivity which results from the small decrement and tends to limit the band-width. It may be said, in fact, that the reduction of noise has been obtained at the expense of decreasing the band-width.

(3) SPACE-CHARGE-LIMITED CONDITIONS

So far we have only mentioned conditions covered by equation (1), which is based on the assumption that the time of transit of the electrons is so small as to have no effect on the noise spectrum, and that the arrival of electrons at the valve anode is " random," i.e. that the emission of an electron is not influenced by the time at which any previous electron was emitted, or arrived at the anode, but is controlled solely by the combination of the external electric field (due to potentials on the various electrodes of the valve) and its own energy of thermal agitation within the cathode. Experimental results support this equation for temperature-limited conditions; but when there is sufficient surplus emission to form a considerable space-charge it is found that the noise is substantially reduced. Indeed, the existence of a finite anode-slope resistance shows that the charge on the anode (which is dependent upon the number of electrons that have previously arrived there) must have some influence on the adventures of electrons emitted at a later time, and therefore casts some superficial doubt on the hypothesis of truly random emission.

---

* See Reference (1).

It is generally agreed that with this space-charge (which is after all the practical working condition in some part of the electrode system of every amplifying valve) the envelope of noise voltage, i.e. the larger fluctuations arising from the random arrivals of numerous electrons, as illustrated in Fig. 1, must produce an opposite fluctuation in anode current, just as much as would occur with any signal voltage developed across the external circuit and appearing as a potential on the anode. (The noise envelope is of course a function of the characteristics of the external anode circuit, as well as of the fluctuation in the current arriving at the anode.) If then the valve has slope resistance R_a and the external circuit is as assumed for equation (1), this effect is believed to cause a reduction in the shot-noise energy by a factor $R_a/(R_a + R)$, giving the Moullin and Ellis equations

$$V_s^2 = \frac{1}{2} I_{av.} e \left(\frac{R_a}{R_a + R}\right) \frac{R}{C} \quad . \quad . \quad . \quad . \quad (4)$$

for the total noise voltage, and

$$V_{df}^2 = 2 I_{av.} e \left(\frac{R_a R}{R_a + R}\right)^2 df \quad . \quad . \quad . \quad (5)$$

for the noise voltage in a narrow band of frequencies df. For lack of a better name, the theory represented by equations (1) to (5) will be referred to as the " pure shot " theory, which does not explicitly consider events taking place within the space-charge, but only the arrival of electrons at the anode. According to this view, shot noise is as a first approximation expressible solely in terms of the magnitude of the current arriving at the anode [equation (1)]; after allowing for the effect of external-circuit characteristics and internal resistance of the valve on the voltage actually generated on the anode [equations (4) and (5)], any corrections required under space-charge conditions are outside the scope of this theory. Actually, however, it is agreed by all experimenters that valve noise is greatly reduced, but never completely abolished, by space-charge limitation of the anode current.

Llewellyn[*] and a number of American workers, on the other hand, take an entirely different view, leading to what may be briefly described as the " thermal noise plus shot noise " theory. They consider that there are two separate sources of noise within the valve. First, random emission from the cathode in the temperature-limited condition must cause an exactly corresponding random arrival at the anode, and hence give rise to " pure shot noise." But Llewellyn considers that when space-charge is present it smooths out the random emission from the cathode, so that arrival at the anode is no longer random, the smoothing being proportional to a factor $(dI/dI_c)^2$; the pure shot noise is to be multiplied by this factor to give the shot noise modified by space-charge. (I is the actual anode current, and I_c the total emission from the cathode.) Thus with complete space-charge limitation, $dI/dI_c = 0$, shot noise should vanish. But in addition to pure shot noise there is said to be " thermal noise " within the valve, for it is argued that since the valve has an effective slope resistance R_a it must on general grounds of thermodynamics give rise to the amount of thermal-agitation noise (or " Johnson noise ") appropriate to such a resistance.

Experimentally, neither theory has proved satisfactory. The difficulty with the thermal-noise-plus-shot-noise theory has been that in order to correspond to the observed values of slope resistance and noise the temperature of the internal resistance must be somewhere about one-half of the cathode temperature; it has in the past seemed a natural assumption that the slope resistance should be at a temperature equal to that of the cathode, but no published paper has covered this question adequately.[*] The value of the total emission I_c for use in the factor dI/dI_c is also a matter of some difficulty, not only with dull emitters but apparently even with pure tungsten cathodes. Pearson, for example, published some experimental results in support of the thermal-noise-plus-shot-noise theory against the pure shot-noise theory.[†] He found it necessary to resort to an extrapolation method to eliminate the effects of external field on the temperature-limited emission from his tungsten cathode; even so, at very small currents (low filament heating) his values of I_c were subject to inaccuracy, for he obtained in the worst case $dI/dI_c = 1{\cdot}03$, i.e. the current to the anode increasing 3 % faster than the reputed total emission from the cathode. The exponents of the pure-shot-noise theory, on the other hand, have not produced any alternative theory for the reduction of shot noise by space-charge.

(4) THERMAL NOISE IN A VALVE

The present author was therefore faced with the problem of first finding the weaknesses and strengths of the two theories referred to above, and then endeavouring to construct a more adequate hypothesis. As regards thermal noise, the author is sufficient of a physicist by training to regard thermodynamics, when properly applied, as an infallible instrument; there appears to be no flaw in Nyquist's thermodynamic proof of the universality of thermal-agitation noise, whatever may be the magnitude of such noise.[‡] Moullin objects that within the valve there is no collision between free electrons and fixed molecules, so that, the mechanism of ordinary resistance being absent, thermal agitation noise which arises in solid conductors should also be absent.[§] But consider a resistance connected between grid and cathode of an amplifying valve; thermal noise arising from this resistance is apparent only as a varying charge on the grid of the valve, and, though we may fairly deduce therefrom that there is a random motion of electrons in the resistance, we need independent evidence to determine how that random motion is caused. Similarly, if the electrons within the anode stream of a valve have random components of velocity, they can produce a " noise " voltage which is indistinguishable from that produced by a metallic resistance, though the mechanism of collision with fixed molecules is not present; we have to rely upon the thermodynamic proof, however, to assure us that the relation between magnitude of noise voltage and magnitude and temperature of resistance is the same in both cases. Actually some correction will be required, since a valve is not a true ohmic resistance but has a curved characteristic. Nyquist's argument is based on the power absorbed, which, when a voltage V is applied to a resistance R, is equal to V^2/R; bearing this in mind, we can find the relation between slope resistance R_a (i.e. tangent to the characteristic) and the value of resistance R' effective for thermal noise in any particular shape of characteristic. For example, if the characteristic is

---

[*] See Reference (2).

[*] A paper on this question by B. J. Thompson and D. O. North was presented at the Rochester meeting of the Institute of Radio Engineers, 16th November, 1936, but so far as the author is aware this has only been printed in abstract form.

[†] See Reference (3).

[‡] See Reference (4). An alternative to Nyquist's derivation of the magnitude and temperature dependence of thermal-agitation noise is given at the end of this paper.

[§] See Reference (1).

$$i = aV^{3/2} \tag{6}$$

the power may be expressed as

$$P = iV = aV^{5/2} \tag{7}$$

To find the additional power absorbed when a small extra voltage, e.g. a " noise " voltage, is superimposed upon the mean applied voltage V, equation (7) is differentiated. Thus

$$\frac{dP}{dV} = \frac{5}{2}aV^{3/2} \tag{8}$$

Now for a true ohmic resistance of value R' we have

$$P = V^2/R'$$

$$dP/dV = 2V/R' \tag{9}$$

We therefore adopt as the definition of the value R' of resistance effective for thermal noise in any circuit element, the expression derived from (9),

$$R' = 2V\bigg/\frac{dP}{dV} \tag{10}$$

For the 3/2-power characteristic, therefore, we find from (8) and (10) that

$$R' = \frac{2V}{(5/2)aV^{3/2}} \tag{11}$$

This is related to the slope resistance R_a by writing

$$\frac{1}{R_a} = \frac{di}{dV} = \frac{3}{2}aV^{\frac{1}{2}} \tag{12}$$

so that combining (11) and (12) we have

$$R' = 6R_a/5 \tag{13}$$

Another important characteristic is the 5/2-power law:—

$$i = aV^{5/2} \tag{14}$$

$$P = aV^{7/2} \tag{15}$$

$$\frac{dP}{dV} = \frac{7}{2}aV^{5/2} \tag{16}$$

$$R' = 2V\bigg/\frac{dP}{dV} = \frac{4}{7}\cdot\frac{1}{aV^{3/2}} \tag{17}$$

But $$1/R_a = \frac{5}{2}aV^{3/2}$$

Therefore $$R' = \frac{10}{7}R_a \tag{18}$$

The third characteristic of practical interest is exponential:—

$$i = ae^{bV} \tag{19}$$

$$1/R_a = abe^{bV}$$

$$P = aVe^{bV}$$

$$dP/dV = ae^{bV}(1 + bV) \tag{20}$$

$$R' = \frac{2bVR_a}{1 + bV} \tag{21}$$

It will be noticed that in all these cases the value of R' is *greater* than R_a, so that the thermal noise from a valve of slope resistance R_a is slightly greater than from an ohmic resistance of value R_a. Alternatively, if the temperature of the valve's internal resistance is calculated from the observed noise voltage and the value of R_a (in place of the correct value R'), the temperature so deduced will be *higher* than the true temperature by a factor such as 6/5 or 10/7.

(5) "PURE SHOT NOISE" THEORY

The next question is whether there is in addition to thermal noise a separate shot noise; the author believes there is not. Thermodynamic reasoning, as usual, indicates the overall relations between the valve, regarded as one unit, and the external circuit, without revealing the internal mechanism within the valve. The random arrival of electrons at the anode, in other words " shot noise," is the expression of the fact that the electrons have a certain random component of velocity which represents the thermal agitation of the electrons, and must therefore be related to thermal noise. It will be shown later in this paper that a proper consideration of the temperature of the valve's internal resistance makes it possible to unify these two aspects of the phenomenon. But, if so, what of the pure shot equation [equation (1) above] for which Moullin states that the sole condition for shot-noise power to be proportional to the magnitude of the mean anode current is that the arrivals of all electrons shall be random? On this view, the noise does not depend upon the magnitudes of the random components of electron velocities, i.e. the temperature of the electrons, but only on the absolute independence of the events constituted by the arrivals of the several electrons at the anode; and the general statement made by some writers that the space-charge " smooths out the irregularities of the emitted current " is unsatisfactory, since it can be shown that space-charge does not destroy the random nature of the electron emission from the virtual cathode.

Moreover, a direct antithesis to the frequency-spectrum method employed in Moullin's derivation of equation (1) is provided by the work of T. C. Fry.* Fry denies the existence of a frequency spectrum of shot noise, and states that " If electrons have been emitted in a statistically steady stream for infinite time past, the probability of the spectrum corresponding to this emission having any pre-assigned ordinate at any given frequency is zero, and the probability that the ordinate exceeds any finite quantity, however large, is unity." This claims to rule out any derivation in terms of a spectrum determined solely by the number of electrons emitted. But examination of the two papers (Fry's and Moullin's) reveals the critical point at which their analyses diverge to such opposite conclusions. Both authors find a Fourier integral representing the pulse produced by the passage of a single electron from cathode to anode, in which the shape of pulse does not affect those frequencies whose period is large compared with the time of transit of the electron, i.e. the frequencies which are of interest for normal radio work; they then proceed to sum the effect of a large number of electrons arriving in random phase, and the crux of the question is the length of period over which the summation is to extend. Fry sums over an *infinite* period, and hence finds that the resultant for any frequency is infinite, being the sum of an infinite number of vectors of equal magnitude but random phase. In corroboration of this result, Fry quotes Einstein's equation for the Brownian motion as a parallel case where the sum of a number of vectors in random phase tends to an infinite resultant as the number of vectors tends to infinity. Moullin, on the contrary, sums over a *finite* period such that N electrons have passed, and quotes a statistical theorem that the sum of N equal vectors in random phase is equal to $N^{\frac{1}{2}}$ times the magnitude of each single constituent vector. The present author believes that the period of summation should be finite and related to the time-constant of the apparatus used to measure the noise voltages; the time-constant in question should probably be that of the first integrating element in the amplifying and measuring chain, the nature of which naturally depends upon the actual apparatus in use. This rules out Fry's result, but does not thereby justify Moullin's: it remains to be determined whether the period of the summation is in fact such that the statistical theorem which he quotes is applicable.

---

* See Reference (5).

There seems to be some doubt among statisticians as to the significance of any limiting value of a property of a collection of events which are unlimited in number. (See, for example, N. Campbell's paper on "The Statistical Theory of Errors,"* particularly pages 802 and 803 of that paper.) The problem is, that if ϵ is the deviation of the actual sum of a particular collection of N unit vectors in random phase, from the statistical value of $N^{\frac{1}{2}}$, how large must N be in order that the probability of exceeding a tolerable limit of error ϵ_0 shall be less than a chosen small quantity δ? This is a complicated question, since it involves firstly the *limit of error* ϵ_0 which can be tolerated, and secondly the *probability* δ of exceeding this limit, which shall for practical purposes be regarded as zero probability. That N can be sufficiently large in atomic phenomena for the difference between actual and statistical values to be imperceptible is evidenced by the successful treatment of ordinary liquids and gases in bulk as continuous fluids. A recent paper by E. N. Rowland† discusses the problem mathematically. Taking the case of random distribution of events governed solely by a constant probability (this corresponds to what was rather loosely described as "constant mean current" in connection with Fig. 1), and supposing observations to be made of the resultant effects of numbers of these events occurring in a series of limited intervals of time, the problem considered by Rowland is whether any meaning can be attached to the mean value of the resultant effect averaged over infinite time, and, if so, whether it can be related to the averages obtained over finite intervals of time. He concludes that, provided the intervals of observation are sufficiently long, we may take the mean value found in any single interval of observation as a reasonable physical estimate of the corresponding quantity averaged over all time. He does not, however, give a criterion of the time which is sufficiently long, nor the corresponding tolerance of error implied by the term "reasonable physical estimate." In the terms in which the problem was stated above, Rowland's paper gives rigorous proof of the experimental deduction from the kinetic theory of fluids, that as N tends to infinity, ϵ_0 and δ tend to zero; but it does not give the numerical relation between the three quantities for values of N other than infinity. The "spectrum" derivation of equation (1) is therefore left as of unknown quantitative accuracy but at least qualitatively correct;* it fails to explain the reduction of noise by space-charge, but it may be that this is due to a misunderstanding of the event constituting the unit vector under conditions of space-charge.

(6) PHYSICAL CONSIDERATIONS

As an introduction to an alternative method of calculation, let us return to Fry's paper, and adopt his definition that "The Schroteffekt will be measured by the difference between the energy actually dissipated in the circuit and the energy that would be dissipated if the electron stream were a uniform flow of a continuous fluid." By calculating the energy that would be dissipated in the attached circuit for a single electronic charge instantaneously transferred from cathode to anode of the valve, he deduces the expression

$$W = \nu \overline{W}_1 + W_0 \quad . \quad . \quad . \quad . \quad (22)$$

where W is the total energy dissipated in the circuit, W_0 the steady-current energy, ν the number of electrons emitted, and $\overline{W}_1$ the mean over all electrons of the energy which would be dissipated in the anode circuit on the emission of a single electron. The noise energy is then $\nu \overline{W}_1$.

Before proceeding further, it is necessary to deal with the doubt, expressed by Fry, whether the disturbance to the external circuit caused by the arrival of a single electron can obey the same laws as the disturbance due to a large charge; for example, how can the charge on a condenser decay exponentially if that charge consists only of a single indivisible electron? This problem was also considered by Rowland, who calculated the shot noise in a circuit connected to a temperature-limited diode according to two different hypotheses: (*a*) the electrons arrive and depart suddenly from the anode, having a random distribution of length of life on the anode, subject only to the half-life constant being such as to give the correct mean value of current; or (*b*) the charge on the anode due to each electron decays exponentially, just as a charge consisting of numerous electrons would. He is surprised to find that it is hypothesis (*b*) which gives the result in accordance with the experimentally verified equation (1). But if one regards the electromagnetic field as the ultimate reality, as in Poynting's theorem for example, this is a reasonable result. The quantum of electromagnetic energy of radio frequencies is so small that the electromagnetic field may here be regarded as continuous. The unit charge remains undivided, but the effective charge on the anode gradually decreases as the electron recedes from it.

---

* See Reference (6).

† *Ibid.*, (7).

* This is, of course, assuming the validity of Rowland's analysis, which the present author is not competent to question.

Returning to equation (22), let us consider the general conservation of energy in the system comprising the valve, anode circuit, d.c. anode supply source, and hot cathode. The d.c. anode supply is not likely to be the source of the noise energy, which is made up of alternating voltages; it might be argued that the random nature of the emission would modulate the steady anode current and thereby enable some of the energy supplied by the battery to be converted to a.c. energy, but one would then expect the noise energy to be proportional to the energy in the steady current. This is not supported by experiment, for the noise energy in a temperature-limited diode depends only on the anode current, and not on the anode voltage; it therefore seems reasonable to suppose that the energy supplied by the battery is always W, the same as would be expended if the current flow were uniform and continuous. The remaining energy $\nu\overline{W_1}$ involves the mean of a number of energies W_1, each of which is characteristic of some individual electron, not of the external circuit or applied voltages; and the most obvious form of energy which is peculiar to every particle of atomic dimensions is thermal energy. It is therefore a plausible hypothesis to suggest that $\nu\overline{W_1}$ is the thermal energy of the electron stream constituting the anode current; it is therefore a function of the temperature of the virtual conductor (or resistance) within the valve. The shot-noise energy is therefore derived from the random component of kinetic energy (i.e. the thermal energy) of the electrons in the anode-current stream, which in turn is derived from the thermal energy of the cathode. Yet despite the obvious suggestion of a "thermal" energy in W_1, Fry finally deduces for the noise energy in a circuit attached to a valve an expression which is a function only of current, and not of any value of energy associated with the electrons constituting the current.

The reason is that he assumed the *instantaneous* passage of each electron from cathode to anode, whereas the only way in which the thermal-agitation energy of the electrons can make itself felt is in variation of the velocity of the electrons between the valve electrodes, a possibility which is excluded by the assumption of instantaneous transit, since this implies that all electrons alike have infinite velocity.

As a purely qualitative example of the effect of transit velocity on the resultant energy expended in the circuit, consider Fig. 2, where C represents the capacitance between the anode and cathode of a valve, and the resistance R the external circuit. If now charge q is transferred from plate "a" to plate "b" of condenser C instantaneously, the energy introduced into C, and afterwards dissipated in R, is $q^2/(2C)$. Now let the charge be transferred from "a" to "b" over a finite time, according to the following rather artificial hypothesis. First a charge $\frac{1}{2}q$ is transferred instantaneously, charging C to a potential $q/(2C)$, then a current flows such as to maintain this potential difference across R until the remaining $\frac{1}{2}q$ has passed. The energy associated with the initial instantaneous transfer of $\frac{1}{2}q$ is $q^2/(8C)$, and will be dissipated at the end of the cycle of operations, when the condenser discharges from potential $q/(2C)$ to zero. The current required in the second stage is $i = q/(2CR)$ [since it is required to maintain a potential $q/(2C)$ across the resistance R] so that the energy dissipated by this current is

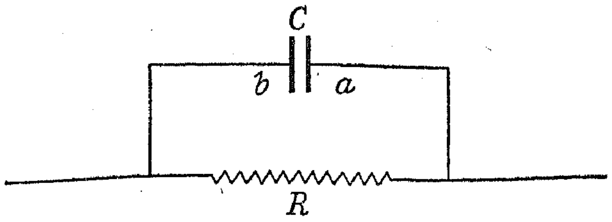

Fig. 2

$$i^2Rt = q^2t/(4C^2R) \quad . \quad . \quad . \quad . \quad (23)$$

where t is the time of current-flow and is given by

$$t = q/(2i) = (q/2)\,(2CR/q) = CR$$

Therefore
$$i^2Rt = q^2/(4C) \quad . \quad . \quad . \quad . \quad (24)$$

The total energy in the cycle of operations is $q^2/(4C) + q^2/(8C) = 3q^2/(8C)$, against $q^2/(2C)$ for instantaneous transit. Thus the electronic velocity of transit should influence the energy dissipated in the external circuit, and the noise energy as calculated by Moullin and Fry on the assumption of instantaneous transit should be a maximum value. This appears to be in general accord with experimental evidence.

It might at first be thought from the above that the reduction of noise in the presence of space charge is due to the increased time of transit of the electrons from cathode to anode; but this is not so, for the time of transit under normal working conditions is short—Moullin (*loc. cit.*) quotes 10^{-9} second—so that the idea of instantaneous transit would be justified on this score. But there is another way of regarding the problem. For example, consider a hypothetical valve in which the electrons are emitted *regularly*, and in such numbers that their periodicity corresponds to a frequency much higher than can be detected by radio apparatus; such a valve would give zero noise, however long or short the transit time of the electrons, for the *regular* wave-form could only produce frequencies above the range of the apparatus, whatever the shape of the constituent pulses of the wave-form. Next imagine a single gas molecule or similar body to be projected across the direction of the current flow, and stop one or two electrons by collision. The stoppage of these electrons will cause a minute pulse to be superimposed upon the steady current-flow in the external circuit, i.e. will give rise to "noise." From this example we see that deviations of electron velocities from a mean value can cause noise, and we must therefore regard the velocities of the electrons in any space-charge-limited valve as made up of two components:—

(*a*) All the electrons have an equal mean velocity, corresponding to the observed steady current, and if they had no other component of velocity the noise would be zero in a practical circuit.

(*b*) Each electron has an individual random velocity of thermal energy, the magnitude of which controls the amount of noise arising in the circuit connected to the valve.

Temperature-limited valves, on the other hand, and any valves whose electrode systems are such that they behave in a similar way, cannot conveniently be regarded from this point of view; for it will be realized from the investigation that follows that it is not possible to assign a temperature to the anode stream of a temperature-limited valve, and in the absence of space-charge there is no significance to be attached to the "mean velocity" shared by all electrons.

(7) TEMPERATURE OF THE ANODE CURRENT STREAM

Earlier in this paper it was stated that a valve which can behave as a resistance must for thermodynamic reasons be a source of noise energy of magnitude appropriate to its resistance and temperature. But it was pointed out that corrections are necessary when the conducting path does not obey Ohm's law, a caveat which applies very strongly, for example, to the temperature-limited diode whose "resistance" is infinite but noise energy finite. It has now been suggested that the source of the noise energy is the random motion of the electrons during their flow from cathode to anode; this is practically identical with the mechanism of thermal noise in metallic conductors.* Since it is thus fair to regard fluctuation noise in a valve having a finite resistance as a thermal effect, it is necessary to find the effective temperature of the electrons constituting the valve's anode current, i.e. the temperature of the valve's resistance.

"Temperature" must first be defined. It is a statistical property of a collection of particles, proportional to the average kinetic energy of random motion possessed by the constituent particles. The term "random motion" is intended to exclude the kinetic energy due to any velocity common to all the constituent particles. Thus a quantity of gas flowing through a pipe at high speed gains additional kinetic energy by its flow, but this does not constitute an increase of temperature, neither does the mean velocity of an electron stream from cathode to anode under the influence of an external source of e.m.f. add to the temperature of the electrons.

O. W. Richardson† showed by thermodynamic reasoning that electrons emitted from and in equilibrium with a hot conductor have a Maxwellian distribution of velocities corresponding to the same temperature as the conductor; he also found this to be in agreement with experimental evidence. It appears that this hypothesis still stands, for Hume Rothery in his book ("The Metallic State," pp. 142–3) states that "within the limits of the experimental methods the electrons emitted from pure metals in a high vacuum have velocities in accordance with the Maxwell law. . . ." It seems certain that if any two systems are capable of exchanging thermal energy, they will in equilibrium be at the same temperature. Such exchange of energy is possible between a hot conductor and electrons surrounding it, by means of both emission and absorption of electrons and reflection at the surface of the conductor.

In a thermionic valve working with a fair amount of space-charge limitation there will be a potential minimum at some point close to the cathode; all space-charge between this and the cathode has a free exchange with the cathode and is therefore at the same temperature. But the passage across the boundary line formed by the potential minimum is an irreversible process: any electron which passes this dividing line is inevitably swept across to the anode. There is therefore no thermal exchange between the cathode and the electrons between potential minimum and anode, and the temperature of these electrons might therefore differ from that of the cathode, and must now be calculated.

From the point of view of noise voltage generated in the external circuit, we are concerned only with electron velocities in the direction of the cathode-anode current stream, and therefore take the components of electron velocity normal to the cathode. This component of the energy of an electron in the outer space-charge* is that with which it left the cathode, less the energy required to pass through the potential minimum, and electrons which are received at the anode are those whose initial velocity component normal to the cathode was greater than was necessary to pass the potential minimum. Richardson (*loc. cit.*) gives the law of distribution of velocities normal to the cathode as

$$N_u du = N \, . \, 2hmue^{-hmu^2}du \quad . \quad . \quad . \quad (25)$$

where N_u is the number of electrons out of a total N which have velocities between u and $(u + du)$, and $1/h = 2kT$. The mean energy of normal components averaged over all electrons leaving the cathode is thus $1/(2h) = kT$. The kinetic energy associated with any group $N_u du$ of electrons is, from (25),

$$W_u = \tfrac{1}{2}mu^2 N_u du = Nhm^2u^3e^{-hmu^2}du \quad . \quad (26)$$

Let u_0 be the velocity which is just sufficient to bring an electron to the potential minimum. Then those electrons which pass over to the anode had initially an aggregate energy W_0 given by

$$W_0 = Nhm^2 \int_{u=u_0}^{u=\infty} u^3 e^{-hmu^2} du \quad . \quad . \quad . \quad (27)$$

---

* The difference is that in a metallic conductor the electrons are in thermal equilibrium with the molecules of the conductor, so that their temperature can be measured by any normal thermometer; in parts of the space-charge, on the other hand, there is no thermal equilibrium between electrons and molecular matter, so their temperature must be calculated from their history since leaving a body having a measurable temperature (i.e. the cathode).

† See Reference (8).

* "Outer space-charge" is a convenient name for space-charge outside the potential minimum.

The number N_0 of electrons passing the barrier is obtained by integrating (25). Thus

$$N_0 = 2hNm\int_{u=u_0}^{u=\infty} ue^{-hmu^2}du \quad . \quad . \quad . \quad (28)$$

$$= Ne^{-hmu_0^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

Evaluation of the integral in (27) gives

$$W_0 = \tfrac{1}{2}Nme^{-hmu_0^2}\left[u_0^2 + 1/(mh)\right] \quad . \quad . \quad (30)$$

and on dividing (30) by (29) the mean initial energy of forward velocity of those electrons which ultimately pass the barrier is

$$\overline{W}_0 = \frac{W_0}{N_0} = \tfrac{1}{2}m\left[u_0^2 + 1/(mh)\right] \quad . \quad . \quad (31)$$

But each electron loses energy $\frac{1}{2}mu_0^2$ in passing the potential minimum, since u_0 was defined as the critical velocity. Therefore the average energy $\overline{W}$ of the electrons when they reach the outer space charge is related to their average initial energy $\overline{W}_0$ by the equation

$$\overline{W} = \overline{W}_0 - \tfrac{1}{2}mu_0^2 \quad . \quad . \quad . \quad . \quad (32)$$

Substituting (31) in (32) now gives

$$\overline{W} = 1/(2h) \quad . \quad . \quad . \quad . \quad . \quad (33)$$

The mean kinetic energy of forward velocity of the electrons is therefore unchanged by passing through the potential minimum, remaining equal to kT. But, considering a metallic conductor as source of thermal noise, it is clear that treatment of an electron stream as a source of comparable thermal-agitation noise requires that the random velocities of the electrons shall be equally distributed both *forward and backward* along the direction of the current-flow. Again, referring to our definition of temperature, we are required to find a mean kinetic energy of random velocity, excluding any drift velocity common to all constituents of the system. We therefore regard the equivalent in volts of the *mean* of the emission velocities of the electrons as being added to the steady anode voltage, and the *deviations from the mean* as the source of thermal-agitation noise. If as an approximation we assume the mean electron velocity to be equal to the mean-square velocity, this means that we simply halve the value of u in all the energy calculations, and divide by four the kinetic energies which depend upon u^2.

The new mean value of kinetic energy of random velocity along the chosen direction is then $kT/4$. But in a perfect gas the mean energy of a particle along any one of a set of three Cartesian axes is equal to the equipartition value $kT/2$; the electrons with mean random energy $kT/4$ have therefore an effective temperature $T/2$, one-half the cathode temperature.

Actually the mean velocity in a Maxwellian distribution is $\sqrt{[8/(3\pi)]}$ times the root-mean-square velocity, so the mean velocity assumed in our approximation is in error by the difference between unity and $0{\cdot}921$, i.e. by $7{\cdot}9$ %. It might at first be thought that this would cause an error of nearly 16 % in the calculated temperature, since this depends upon the square of velocity; but actually it is a question of deducting velocity from one group of electrons and adding to another, so that the error depends upon a difference of squares only, and is small, provided the two squares are nearly equal, which is true in this case.

We have thus shown that in a valve having a resistance which is finite and a function of space charge (i.e. a space-charge-limited valve) the total fluctuation noise is equal to the thermal-agitation noise in a corresponding resistance at a temperature of approximately half the cathode temperature, subject to the following corrections:—

(1) In practical valves, end-effects cause a fraction of the current to be temperature-limited, and therefore probably more "noisy." It might be worth while to try experiments with guard rings, so as to use only the central part of the cathode where the current can be fully space-charge-limited.

(2) The relation between slope resistance and effective resistance for thermal noise must be taken into account. [See, for example, equations (13), (18), (21), which indicate that the effective resistance is normally greater than the slope resistance.]

(3) Although the simple theory of thermionic emission has been well tested for pure metals, there is a slight possibility (though no actual evidence) that with complex emitting surfaces the space-charge adjacent to the cathode might have a temperature different from that of the cathode, though probably retaining a Maxwellian distribution of velocities. This can only occur if quantum effects prevent the free exchange of energy between the space-charge and the interior of the cathode.

(8) MECHANISM OF THERMAL-AGITATION NOISE IN SOLID CONDUCTORS

We have found that the fluctuation noise in a space-charge-limited valve may be represented as a thermal-agitation noise. Conversely, thermal-agitation noise in a metallic conductor may be represented as "shot noise" within the crystal lattice of the material, and without space-charge limitation.

Let us consider a conductor of length Δl and sectional area A, having n_0 free electrons per unit volume, and subjected to a uniform longitudinal electric field of strength E. Within a solid conductor, collisions between electrons and molecules are extremely frequent, and the drift velocity acquired under an applied electric field is small compared with thermal-agitation velocity; we may therefore assume a fixed time of flight τ between collisions, depending solely upon the distance between molecules and the temperature. Denoting the electronic charge and mass by e and m respectively, the mean drift velocity $\bar{u}$ is given by

$$\bar{u} = eE\tau/(2m) \quad . \quad . \quad . \quad . \quad . \quad (34)$$

From the definition of electric current as $i = dQ/dt$, it follows that if there is a small finite variation of charge ΔQ in a time Δt, the average current flowing is $\bar{\imath} = \Delta Q/\Delta t$. If, further, the charge is transferred from one end to the other of a length Δx (i.e. the current flows through a circuit of length Δx) during the time Δt, we may write

$$\bar{\imath}\Delta x = \frac{\Delta Q}{\Delta t}\Delta x = \frac{\Delta x}{\Delta t}\Delta Q = \bar{u}\Delta Q \quad . \quad . \quad (35)$$

since the ratio $\Delta x/\Delta t$ is the mean velocity. Now in the metallic conductor, ΔQ is equal to the product of the charge on an electron and the total number of electrons taking part, Δx may be replaced by Δl (provided the latter is taken small enough), and $\bar{u}$ is given by equation (34), so that

$$\bar{i}\Delta l = (A\Delta l n_0 e)\left[eE\tau/(2m)\right] \quad . \quad . \quad . \quad (36)$$

But the potential difference V between the ends of the conductor is the product $E\Delta l$ of field-strength and length, and the resistance R is the ratio of potential difference to current. Thus

$$R = \frac{V}{\bar{i}} = \frac{E\Delta l}{(An_0e)\left[eE\tau/(2m)\right]} = \frac{2m\Delta l}{An_0e^2\tau} \quad . \quad (37)$$

Comparing the solid-conductor problem with the shot-noise problem in a thermionic valve, we may say that equation (37) has related the mean current to the atomic mechanism, but it remains to determine the fluctuation noise due to the discrete nature of the mechanism. Provided that a current pulse is sufficiently short in duration compared with the period of frequencies capable of being observed, it is possible to evaluate the observable components of the Fourier analysis of the pulse without knowing its shape. For example, in dealing with shot noise in a temperature-limited diode, Moullin and Ellis (*loc. cit.*) showed that, knowing only the integral characteristic of the pulse due to a single electron, $\int i dt = e$, the components of radio frequency can be calculated. In our present problem of the solid conductor the unit pulse is not a current, but is of the form $i\Delta l$, which will be termed a "current-element." Combining expressions (34) and (35) for a single electron, and integrating over its time of flight, we have

$$\int_0^\tau (i\Delta l)dt = e\bar{u}\tau \quad . \quad . \quad . \quad . \quad (38)$$

If, therefore, we replace e in the analysis of Moullin and Ellis by $e\bar{u}\tau$, the components of the Fourier analysis will represent components of $i\Delta l$ in place of i. The time of flight of an electron within a conductor is even shorter than the transit time through a valve, so that the assumption that the pulse is so short that its shape has negligible effect on the radio-frequency components is a fortiori applicable in this case. If $2\pi/p$ is the period of the Fourier series, assumed to be long compared with the duration of the pulse but short compared with radio frequencies, the pulse $e\bar{u}\tau$ is found to be equivalent to a series of current-elements

$$\Sigma i\Delta l = (e\bar{u}\tau p/\pi)(\tfrac{1}{2} + \Sigma \cos npt)$$

In the present problem, unlike the shot effect, $\bar{u}$ may be either positive or negative, so that the constant term (steady current) will vanish on summing over a large number of pulses; the noise current-elements are left as

$$\Sigma i\Delta l = (e\,|\,\bar{u}\,|\,\tau p/\pi)(\Sigma \pm \cos npt) \quad . \quad . \quad (39)$$

Since the time $2\pi/p$ was taken to be very long, the order of harmonic n is high, and we may replace the summation by an integral. The resultant of the noise current-elements in a group of frequencies between n_1 and $(n_1 + dn)$ is then given by the equation

$$(i\Delta l)_{dn} = (e\,|\,\bar{u}\,|\,\tau p/\pi)\int_{n=n_1}^{n=n_1+dn} \pm \cos npt.dn$$

Expressing this in terms of frequency, we have $np = 2\pi f$, $dn = (2\pi/p)df$,

$$\therefore \quad (i\Delta l)_{df} = 2e\,|\,\bar{u}\,|\,\tau\int_{f=f_1}^{f=f_1+df} \pm \cos 2\pi ft.df \quad . \quad . \quad (40)$$

$$(i\Delta l)^2_{df} = 4e^2\bar{u}^2\tau^2\int_{f=f_1}^{f=f_1+df} \cos^2 2\pi ft.df$$

$$= 2e^2u^2\tau^2\int_{f=f_1}^{f=f_1+df} (1 + \cos 4\pi ft)df$$

But on averaging over a considerable period we have $\int \cos 4\pi ft.df = 0$.

The mean-square value of the noise current-elements from a single pulse is therefore

$$(\bar{i}\Delta l)^2_{df} = 2e^2\bar{u}^2\tau^2.df \quad . \quad . \quad . \quad (41)$$

But if there are N electrons taking part, and each electron collides $1/\tau$ times per second, i.e. makes $1/\tau$ separate journeys per second, there are altogether N/τ randomly-phased pulses per second. Since they are in random phase we may add the squares of the random current-elements (this is assuming that the resultant of N random-phased vectors is $N^{\frac{1}{2}}$ times the unit vector), so that the resultant of the noise current-elements in frequency band df has a mean-square value

$$(\bar{I}\Delta l)^2_{df} = 2Ne^2u^2\tau.df$$

and

$$\bar{I}^2 = (2Ne^2u^2\tau.df)/(\Delta l)^2 \quad . \quad . \quad . \quad (42)$$

But if current $\bar{I}$ flows through resistance R the corresponding potential difference is, by definition of R, equal to $\bar{I}R$; the mean-square noise voltage corresponding to (42) is therefore

$$\overline{V}^2_{df} = R^2\bar{I}^2_{df} = (2R^2Ne^2u^2\tau.df)/(\Delta l)^2 \quad . \quad . \quad (43)$$

Using equation (37) to eliminate N, e, and τ,

$$\overline{V}^2_{df} = 2R.2m\bar{u}^2.df \quad . \quad . \quad . \quad . \quad (44)$$

It is clear that the only velocity which will contribute to the current between two points is the velocity along the direction of flow; in other words, $\bar{u}^2$ is to be taken as the mean-square of the component of thermal-agitation velocity in one specified direction. But the equipartition value of the thermal energy contributed by velocity components in one given direction is $\frac{1}{2}m\bar{u}^2 = \frac{1}{2}kT$ for each particle. Equation (44) therefore becomes

$$\overline{V}^2_{df} = 4RkT.df \quad . \quad . \quad . \quad . \quad (45)$$

which is the well-known expression deduced thermodynamically by Nyquist.

Whereas in Nyquist's derivation the internal mechanism was not involved, owing to the use of overall energy-exchanges, in the treatment given above the mechanism is exposed, but all factors peculiar to the material, namely the number of electrons taking part in conduction, time of flight, and mass and charge of the electron, have been eliminated in terms of the resistance and the function kT. Equally with the thermodynamic proof, therefore, the derivation presented above indicates that the thermal-agitation noise is a function only of ohmic resistance and temperature, and not of the structure of the conductor involved. The thermionic valve is therefore included, so long as it has a determinable ohmic resistance and temperature.

It is interesting to note that in Nyquist's derivation there was no upper limit to the frequency to which the expression for noise energy was applicable; something in the nature of quantum restrictions was necessary to prevent the total energy from becoming infinite if the frequency range was extended to infinity instead of being confined to the radio band. In the derivation used above it is obvious that the expression ceases to be valid when the frequency has a period comparable with the time of flight of the electron within the conductor, for the Fourier analysis must then be modified and will depend upon the shape of the pulse. There is therefore seen to be a limit to the validity of the expression, though the limiting frequency will be somewhere in the region of heat radiation.

(9) APPLICATION TO THE THERMIONIC VALVE

The remaining difficulty is the transition from a temperature-limited to a space-charge-limited regime in a thermionic valve: we have yet to decide what constitutes the essential distinction between the two states. It would seem that the presence of a potential minimum, however small the barrier which it imposes, is one criterion; for, provided there is such a barrier, some electrons emitted will return to the cathode, and the space-charge adjacent to the cathode will be in thermal equilibrium with it. The initial temperature of the space-charge will then be fixed at the cathode temperature, and, as we found above, the effective temperature of the outer space-charge is about half that of the inner space-charge. The transition in any real valve will be gradual, owing to lack of uniformity both of the anode-to-cathode field along the length of the cathode and of the cathode temperature.

Another criterion of the state in which the thermal-agitation treatment is applicable is that the field from the anode should terminate on space-charge, not on an actual metallic electrode. For if the field from the anode ends on a metallic electrode, the emergence of any electron from that electrode constitutes a disturbance; but if the field terminates on electrons, constituting space-charge, which are moving towards the anode with approximately uniform velocity, the presence of electrons travelling at the exact mean velocity at every point does not create a disturbance; it is only the deviations from the mean which are effective.

In a diode, but not in more complex valves, the two criteria are identical. As an example of the more complex valves, consider a screen-grid tetrode. In this the screen-anode space corresponds nearly to a temperature-limited diode with the screen as virtual cathode. Most of the anode field terminates on the screen wires, so that the injection of electrons into this space through the screen may be expected to produce a shot noise at the anode, of the magnitude predicted by equation (1). The same effect is to be expected to a less extent in triodes, depending upon the closeness of the grid winding; this may explain the experimental results of F. C. Williams with an L.S.5 triode.*

(10) TWO RESISTANCES IN PARALLEL

Another point which is of importance for the comparison of any theory of thermal noise with experimental results is the magnitude of the resultant fluctuation voltage from two resistances at different temperatures connected in parallel. Transferring equation (45) from voltage back to fluctuation current, which we saw was the fundamental phenomenon, we find that

$$\bar{I}^2_{df} = (4kT/R)\,.\,df$$

where $\bar{I}^2$ is the mean-square fluctuation current. But $\bar{I}^2$ is made up of a very large number of random vectors, so that to combine two such currents, arising in resistances R_1 and R_2, we merely add their squares. Thus

$$\bar{I}^2_0 = \bar{I}^2_1 + \bar{I}^2_2 = (4kT_1/R_1 + 4kT_2/R_2)\,.\,df \quad . \quad (46)$$

Equation (46) for the total fluctuation current flowing when two resistances R_1 and R_2 at temperatures T_1 and T_2 are connected in parallel reduces to

$$\bar{I}^2_0 = 4k \cdot \frac{R_2T_1 + R_1T_2}{R_1R_2} \cdot df \quad . \quad . \quad . \quad (47)$$

But the resultant resistance of the two in parallel is $R_0 = R_1R_2/(R_1 + R_2)$. The noise voltage corresponding to the current given in equation (47) is therefore

$$\bar{V}^2_0 = R^2_0\bar{I}^2_0 = 4k \cdot \frac{R_1R_2(R_2T_1 + R_1T_2)}{(R_1 + R_2)^2} \cdot df \quad . \quad (48)$$

This equation has been previously derived by Llewellyn.†

(11) ACKNOWLEDGMENTS

Any scientific worker is indebted to his predecessors, but the present author, owing to his own lack of experimental work on the problem, is particularly indebted to the writers of previous publications, even though he has disagreed with them. In particular, his thanks are due to Mr. E. B. Moullin, who first interested him in this subject and with whom he has had a number of valuable discussions.

The author is also indebted to Marconi's Wireless Telegraph Co., Ltd., for the preparation of lantern slides used in the presentation of this paper.

REFERENCES

(1) E. B. Moullin and H. D. M. Ellis: *Journal I.E.E.*, 1934, vol. 74, p. 323.
(2) F. B. Llewellyn: *Proceedings of the Institute of Radio Engineers*, 1930, vol. 18, p. 243.
(3) G. L. Pearson: *Physics*, 1935, vol. 6, p. 6.
(4) H. Nyquist: *Physical Review*, 1928, vol. 32, p. 110.

---

* See Reference (9).

† *Ibid.*, (10).

(5) T. C. Fry: *Journal of the Franklin Institute*, 1925, vol. 199, p. 203.

(6) N. Campbell: *Proceedings of the Physical Society*, 1935, vol. 47, p. 800.

(7) E. N. Rowland: *Proceedings of the Cambridge Philosophical Society*, 1936, vol. 32, p. 580.

(8) O. W. Richardson: " Emission of Electricity from Hot Bodies," 2nd edition, London, 1921.

(9) F. C. Williams: *Journal I.E.E.*, 1936, vol. 79, p. 349.

(10) F. B. Llewellyn: *Proceedings of the Institute of Radio Engineers*, 1930, vol. 18, p. 256.

DISCUSSION BEFORE THE WIRELESS SECTION, 5TH JANUARY, 1938

Mr. E. B. Moullin: In his discussion of Fig. 1 the author points out that the deviations from the mean increase with the current, even though the flow of electrons becomes relatively smoother. It is a basic theorem of statistics that, in circumstances such as those depicted in Fig. 1, the mean-square departure from the mean increases in direct proportion to the mean value. It is both well known and remarkable that this theorem, true in the limit of very large numbers, is substantially true when the number of events is quite small. This can be illustrated well by means of the diagrams in Fig. 1. Thus we find that the mean-square departures from the mean in Cases A and D of Fig. 1 are 1·3 and 0·6 respectively, if the baseline is taken as 20 units of time. There is no reason why we should take A, B, C, or D, in preference; and therefore the natural course is to take these possible chances of succession one after the other, and then the long-time mean of the mean-square fluctuation is 0·92. According to the statistical theorem I have mentioned, if we have 5 times the current the mean-square departure should be 5 times this amount, i.e. 4·6. We have a particular case of this in the sum curve of Fig. 1; and the mean-square departure is 6·3. Thus we should expect the value 4·6, whereas we get 6·3; but this is too rough an approach to be a fair test. There are 10 possible summation curves of Cases A to E taken in pairs, 10 more possible combinations taken three at a time, and 5 combinations taken four at a time. I have made the 10 possible additions two at a time, and find the mean-square deviation of these comes to 2·27. For the 5 possible combinations four at a time I find the mean-square deviation is 4·53. As an example, let us suppose that the mean-square deviation is proportional to the mean, and equal to 4·53 when the mean is 4. Then, when the mean current is 1, 2, 4, or 5 electrons (I have not worked out the figures for 3), the deviations should be 1·14, 2·27, 4·53, and 5·67 respectively. The measured deviations for the sample curves given in Fig. 1 are in fact found to be 0·92, 2·27, 4·53, and 6·3. The agreement between these sets of values is fairly good, and the example serves to illustrate this important theorem.

I feel that Nyquist's theorem still requires to be stated more rigidly. In its present form it is certainly correct for all linear networks which do not include thermionic valves, but when such are included we still do not know how to apply it to give the correct result, though I consider the author's method of obtaining the effective value of the internal resistance (page 100) is a notable step forward to this end.

Section (5) begins by stating the author's belief that shot and thermal effects are different aspects of the same general principle. I have always thought, and I think often stated, that until we had a general theorem which would include shot and thermal effects as special cases of one general principle, this very intricate noise problem would remain in an unsatisfactory state. In my opinion Section (8) of this paper is a real step towards this general treatment, and is perhaps the most illuminating theoretical contribution to the whole subject which has been made in the last decade. I have often tried to obtain the thermal-agitation formula from the shot-effect principle, and have always failed dismally.

Lately I was talking to Prof. Schottky in Berlin, and he told me he had derived a formula for the thermal-agitation principle from a shot-voltage mechanism. The proof has not yet been published, and I shall be interested to see whether it is the same as the present author's derivation.

From one point of view it has always been difficult to understand why the passage of a steady current through a resistance did not increase the fluctuation voltage produced by it. For if most methods of calculating the shot voltage produced by a temperature-limited thermionic current (for example, that which Mr. Ellis and I produced some years ago),* are examined critically, it is difficult to find any step in the argument which could not be applied to effects inside a conductor. I think the paradox has been resolved by the analysis of Section (8) of this paper: for this brings out clearly that the number of random events in the conductor are governed by the number of free electrons in it, and not by the average current passing through it. Such a current will not produce a first-order effect on the number of random events, and will possibly produce no effect at all.

In spite of all the work that has been done, however, the background noise in an amplifier valve still cannot be calculated. In the last 6 months Prof. Schottky has produced several long papers on the shot effect in space-charge-limited conditions. I have so far been unable to disentangle fully his basic idea from the analysis in which it is clothed. Schottky's formula is appreciably better than the old classic expression of Llewellyn, but it is still in bad agreement with facts.

On page 104 the author observes: " In practical valves, end effects cause a fraction of the current to be temperature-limited, and therefore probably more ' noisy.' It might be worth while to try experiments with guard rings, so as to use only the central part of the cathode where the current can be fully space-charge-limited." That work has just been completed in Germany,† but the results do not appear to approach the theoretical value of Schottky more closely than they would have done if no precautions had been taken.

Dr. F. C. Williams: The author deals with this subject from the point of view of the physicist, whereas I have examined it experimentally.

---

* *Journal I.E.E.*, 1934, vol. 74, p. 323.

† H. Jacoby and L. Kirchgessner: *Wissenschaftliche Veroffentlichungen aus den Siemens-Werken*, 1937, vol. 16, p. 42.

I do not feel competent to question the validity of the author's application of a thermodynamic principle to the problem of thermionic balance, but I see several difficulties in applications of this sort which have been made in the past. Thus Llewellyn considered a box containing a thermionic valve and a resistance uniformly heated to a temperature τ sufficient to enable a copious emission of electrons to take place. On that hypothetical experiment he based his result that the noise in the valve must be a thermal noise and must be given by $\overline{V}^2 = 4\rho k\tau df$. Experiment persists in showing, however, that the relevant value of τ would not be the temperature of the box but half this value. Thus there is either some flaw in Llewellyn's original analysis or some factor which militates against the truth of such applications. Further, it may be noted that in this uniform-temperature box we apparently have a steady transfer of energy from the resistance to the valve, which is a contradiction of the thermodynamic law. If, now, a grid is introduced into the valve and a resistance connected between the grid and cathode, the fluctuations generated in the valve are vastly augmented, and there would probably be a transfer of energy from the valve to the resistance. My own experiments show that even if the resistance is not present but the grid is simply short-circuited to the cathode, the apparent temperature of the anode stream may be from 10τ to 40τ. Until these difficulties have been satisfactorily resolved, some doubt must attach to a thermal representation of valve fluctuations based on thermodynamic considerations.

Thus if valve noise be interpreted in terms of thermal agitation, it is often necessary to assume an effective temperature very much higher than the melting point of the cathode material; further, the temperature is a function of the electrode structure and of the operating conditions. Such interpretation appears rather artificial, and has little to offer which counterbalances the enormous technological advantages of the shot interpretation.

Mr. Moullin mentioned the work of Schottky in attempting a shot explanation of the fluctuations in space-charge-limited valves. Schottky has produced a curve calculated on that basis which Mr. Moullin says is little better than earlier attempts as regards its agreement with experiment. The curve is, however, not vastly out of agreement with some experimental results—the discrepancy being of the order of 1·5 to 2. On the other hand, the only experimental verification of the thermal expression has been made either in the retarding-field region or very close to it, where the $\tau/2$ discrepancy from Llewellyn's original formula is found. It is interesting to note that this discrepancy can readily be explained in terms of pure shot theory.

In conclusion, I may perhaps mention that the process used by the author in Section (10) to derive the fluctuation generated across two resistances in parallel was employed in one of my own papers* to deduce the fluctuation voltage generated by a space-charge-limited current traversing a resistance. It was noted then that the method could be applied to the analogous problem of two resistances in parallel.

Mr. O. E. Keall: The published results of shot-noise investigations carried out in this country and also in Japan and Germany 2–3 years ago show a considerable discrepancy both in the temperature-limited region and also in the space-charge region, although the valves concerned had similar anode-current/heater-current characteristics. It would therefore seem that an experimental investigation of this subject on a much larger scale than has hitherto been conducted is desirable. Particularly is this the case because hitherto the investigations have been limited to consideration of one or two of the variables in the valve only (either resistance or current). In spite of the elegance of the author's analysis, I do not see how one can apply the formula he gives in the case of the valve with infinite resistance or even in the case of the valve with negative resistance, because experimental results indicate that there is no great difference between these cases and that of the valve with normal characteristic.

Mr. W. H. Aldous: A theory of fluctuation noise should agree with practice over at least part of the range of the valve characteristic, and I do not think it can be said to be a complete theory until it agrees with practice over the retarding-field region, the normal working region, and the temperature-limited region. I do not know in which of these regions the author's theory agrees with experiment.

Dealing with Section (4), I do not see that Nyquist proves the universality of thermal-agitation noise by his treatment. He merely says that some source passes energy into a resistance outside itself. Now the second law of thermodynamics does not say what that emerging energy should be, and why one should assume it is thermal-agitation noise. The author proposes to use a value of resistance in the thermal-agitation formula different from the valve slope resistance, which he derives from consideration of the power absorbed, given by $P = V^2/R'$. For a characteristic of the form $i = aV^n$ this gives

$$R' = \frac{2n}{n+1}R_a$$

If he had started instead with

$$P = I^2R''$$

(which is actually the form used by Nyquist), and therefore

$$R'' = \frac{1}{2I}\frac{dP}{dI}$$

a different value of effective resistance would have been obtained, namely

$$R'' = \frac{n+1}{2}R_a$$

I think the idea that the power is produced in the valve itself is incorrect. It arises from the random transit of the electrons across the electrode space producing a fluctuating voltage across the load. Since the valve is in parallel with the load, a certain proportion of the power will be dissipated in the valve slope resistance, which therefore requires no modification. This is in accord with the experimental work of both Schottky and F. C. Williams, who have shown that, in the retarding-field region represented by equation (19), the shot noise is given accurately by equations (4) and (5). This can be seen to be true on general grounds, since in the retarding-field region there is no question of space-charge control, all electrons with initial energy greater than some minimum reaching the anode. These electrons will arrive randomly in time, and it is this randomness in time—not randomness in velocity—which gives rise to the shot voltage across the load.

---

* *Journal I.E.E.*, 1936, vol. 79, pp. 352 and 354.

At the end of Section (7) the author tries to show that a deviation of velocity from the mean can cause noise; but I think he takes an unfortunate example. A few of the electrons are stopped by positive ions, and this means that the steady rate of arrival of the electrons at the anode is altered, a condition which must give rise to shot noise. With regard to the author's method of deriving the value $\tau/2$, since in the case he considers there are so many more electrons with low forward velocities than with high forward velocities, the deviations about the mean will not be randomly distributed. Therefore I feel very doubtful as to whether the mean value of kinetic energy of random distribution along the chosen direction can be taken to be $k\tau/4$.

Finally, I should like to put forward a plea for the use of conductances rather than resistances in these fluctuation formulae, in view of the consequent simplification of the mathematics. Thus equation (48), if extended to several conductances $S_1, S_2, \ldots S_n$ in parallel at temperatures $\tau_1, \tau_2, \ldots \tau_n$ respectively, becomes simply

$$\overline{V}_0^2 = 4k\frac{\sum_1^n S_n\tau_n}{\left[\sum_1^n S_n\right]^2}df$$

where each conductance is now associated with its own temperature.

Dr. W. F. Rawlinson: I should like to ask the author a question in relation to the assumed or calculated temperature of the electron stream. Although the valve has been evacuated it must still contain a large number of molecules or atoms. Since the author derives for the electron-stream temperature a value which is about half the cathode temperature and which bears no relation to the anode temperature, and since his derivation depends on the equipartition theory, are we to assume that the residual gas between the anode and cathode is also at half the temperature of the cathode?

Considering a cathode which gives perfectly uniform emission and therefore no shot noise, and, according to the author, no thermal-agitation noise, will the electron stream be at absolute zero temperature; and, if so, what is the temperature of the residual gas?

Mr. H. D. McD. Ellis: The paper is a valuable contribution towards the solution of these fluctuation problems; seeing the elegance of the author's treatment I regret that I was not trained as a physicist also. I am particularly struck by the way in which the author has cleared up some of the troubles with which Mr. Moullin and I were concerned when we were preparing our paper[*] about 5 years ago. For example, I am interested in his methods of obtaining the thermal-agitation voltages in two resistances at different temperatures, and of determining the effective anode impedance of a thermionic valve from power considerations. We obtained the correct solution to the first of these problems 5 years ago, but not, I fear, in so neat a fashion.

Mr. W. S. Percival (*communicated*): On page 100 the author introduces a correction to the slope resistance of a diode valve to take account of the curved characteristic for purposes of noise calculation. I am unable to follow the author's argument, but equation (21)[*] leads to a result opposed to well-known theory. If in this equation we make V negative and large compared with $1/b$, i.e. if we work in the retardation region with a small current, so that we can neglect the space charge, then approximately, $R' = 2R_a$. It has, however, been shown by F. C. Williams[†] that this is the one condition in which the diode must behave as an apparent (slope) resistance at half the cathode temperature without any correction being necessary.

A more satisfactory line of argument is to consider a choke placed in the anode circuit of the diode, the anode of which is connected via a condenser to a transmission line terminated by a resistance equal to the slope resistance R_a of the diode. Under these conditions the noise power transferred from the resistance to the diode and hence from the diode to the resistance, depends simply on the slope resistance of the diode and is independent of the curvature of the characteristic. If the diode could be at the same temperature throughout and there were no external source of energy, then Nyquist's original argument[‡] would hold. In fact, however, the cathode of the diode is at one temperature, its anode is at another, and external energy is available from a battery. The original proof of Nyquist does not, therefore, apply.

On page 103 the author states " There is therefore no thermal exchange between the cathode and the electrons between potential minimum and anode." This cannot be true, however, since the electrons between the potential minimum and the anode form part of the space charge. Let us suppose that there is a certain rush of electrons past the space-charge minimum, then that portion of the space charge will be increased. The increased negative field will prevent some electrons which would otherwise pass the minimum from doing so. Some of these electrons will fall back into the cathode. Hence those electrons which have passed the potential minimum can still, indirectly, influence the cathode. It is true, as the author says, that no electrons which have passed this point can return. Their electric field is still effective, however, and can influence those electrons which are closer to the cathode.

Prof. W. Schottky (Germany) (*communicated*): Contrary to the " relative temperature theory," in which the electron stream is compared to a gas having a uniform velocity superimposed on the Maxwellian motions of the molecules, I consider that there is no moving co-ordinate system with respect to which the velocity distribution of the electrons has an effectively Maxwellian character: the indefinitely large velocities in a backward direction are missing.

---

[*] *Journal I.E.E.*, 1934, vol. 74, p. 323.

[*] The a in this formula appears to be a misprint [corrected for the *Journal*].

[†] *Journal I.E.E.*, 1936, vol. 78, p. 326.

[‡] *Physical Review*, 1928, vol. 32, p. 110.

Our work shows that the rigorous treatment of the space-charge problem appears rather as a shot effect which approaches the $\tau/2$ law when the current is sufficiently reduced by space charge. I have explained how in a retarding field (without space charge) it is possible to find the "thermal complement" of the cathode-to-anode electron stream, which would make the whole system equivalent to a thermal one. I have not, however, succeeded in finding the "thermal complement" for a space-charge current.

An objection to the "relative temperature theory" is that the rigorous theory, under ideal conditions, yields the value $1\cdot272\tau/2$ for the apparent temperature.* This departure from the $\tau/2$ law is large compared with that which can arise from the insufficient consideration of mean drift velocity suggested towards the end of Section (7) of the paper.

I have nothing to add to Section (8), which is more complete than my own work on this point, since it gives the constant, $4k$, correctly. But the following may be of interest. For the region of very high frequencies, I have calculated the frequency spectrum on the assumption that all electrons have equal velocities in the direction of the current under investigation, but that the lengths of free path are distributed in a random manner. This can be compared to the flicker effect, by considering that the mean impulse time of an electron in the conductor corresponds to the mean life of an extraneous molecule in the flicker effect. One then calculates the energy which on short-circuit is dissipated by this frequency spectrum in an ohmic resistance, and finds (except for a factor approximately equal to unity, of which I cannot yet be certain), the value $k\tau/t$ where t is the mean impulse time. This result is very satisfactory, for it shows that the thermal energy of a degree of freedom in an electronic conductor will on the average be absorbed and restored as often as corresponds to the mean impulse time of electrons, irrespective of the number of electrons on the conductor.

Mr. D. A. Bell (*in reply*): Both Mr. Moullin and Mr. Aldous have doubts as to Nyquist's theorem. Undoubtedly, care is necessary in applying it; it was for this reason that a substantial part of the paper was devoted to discussion of the values of "resistance" and "temperature" to be used. I would go so far as to say, however, that the difficulty is not, as Mr. Moullin suggests, that we do not know how to apply Nyquist's theorem to thermionic valves in general, but that (*a*) we cannot make a valve conforming to ideal laws, and (*b*) we have not yet seen how to modify the theory for valves which are only partially space-charge-limited.

Much confusion seems to arise from the idea of some absolute distinction between different types of fluctuation noise; thus Mr. Aldous states "Now the second law of thermodynamics does not say what that energy should be, and why one should assume it is thermal agitation noise." In my view there is one *assumption*, and one only, that is made in setting out Nyquist's theorem; and that is that it is possible to have two elements of an electric circuit which are connected by electrical conductors, but isolated from each other so as to prevent the direct exchange of other forms of energy. I do not think this is any more unreasonable than the ideas used throughout classical thermodynamics, such as the perfect gas, perfect heat insulator, or perfectly-reversible heat engine. Given this assumption, when a space-charge-limited valve and a resistance are connected together, there must be energy transferred from the valve which is proportional to the temperature of the valve's resistance, and surely such energy must be described as thermal-agitation energy, whatever the internal mechanism by which it is produced.

With regard to the experiments of Prof. Schottky and his colleagues, which Mr. Moullin mentions as having full precautions in the way of guard-rings, is not the best criterion of achievement of the "ideal" system the anode-current/anode-voltage law? In Prof. Schottky's diode one would expect a 3/2 power law, but the published curves seem to indicate a law nearer 1·3 than 1·5 for tube No. 6; tube No. 7 is sufficiently complex to have other difficulties.

If Dr. Williams is not prepared to discuss my analysis, it is doubtful whether there is much value in replying to his objections to Llewellyn's work. As will be seen from a study of Section (7) of the paper, I believe the resistance of the valve must be at a temperature different from that of the cathode, so the idea of a uniform-temperature enclosure including the whole of a valve is ruled out. The same section gives the basis for my belief that there is real significance in a modified $T/2$ law for a space-charge-limited valve. In the case of the triode amplifying valve, with a resistance connected between grid and filament, I see no reason to depart from the normal theory of the amplifying valve, in which the output a.c. power is derived from the source of steady anode potential, even if that source be only the mean initial velocity with which the electrons are emitted from the cathode. As explained at the end of Section (9) of the paper, I do not consider Dr. Williams's results with grid short-circuited to cathode to be incompatible with my theory. The shot interpretation is only applied to space-charge-limited valves by the addition of an arbitrary correction factor (or "smoothing factor") which Dr. Williams denotes by A; until this factor has been given a rational basis it cannot be regarded as having any place in a scientific theory, or used (other than empirically) as an aid to the design of apparatus.

In his statement that in the retarding-field region the $T/2$ law can be readily explained in terms of pure shot theory, Dr. Williams is presumably referring to a paper in which he examined this question.* I regret that I cannot accept his treatment as exact, though it gives a useful approximation. For he assumes that $Ve>>kT$, and $x>>kT$, where V is the working anode voltage, ranging from about $-0\cdot5$ to zero, T the cathode temperature, and x the work function of the cathode. Using values given by Benjamin, Cosgrave, and Warren in a recent paper,† I have prepared the figures in Table A for various cathodes.

Thus for an oxide-coated cathode (which was the type used by Dr. Williams for the retarding-field experiment) x is about 10 times kT, but Ve is only about 6 times kT at the extreme limit of the range ($V = -0\cdot5$ volt); this does not seem to justify the approximations adopted for the whole range 0 to $-0\cdot5$ volt. I do not think the thermal theory has previously been set out in sufficient detail for application to practical valves [see Section (3) of the paper on previous versions of the thermal theory].

---

* *Wissenschaftliche Veroffentlichungen aus den Siemens-Werken*, 1937, vol. 14, p. 15.

* *Journal I.E.E.*, 1936, vol. 78, p. 326.

† *Ibid.*, 1937, vol. 80, p. 401.

I should like to support Mr. Keall's suggestion for more comprehensive experimental work, particularly as I personally have no facilities for experiment and must rely upon the published results of others to verify my hypotheses.

All valves which have an infinite resistance are, in my opinion, operating under conditions which correspond to temperature-limitation, whatever may be the space-charge distribution at some point remote from the anode; the " pure shot noise " formulae are therefore applicable to such valves.

Mr. Aldous mentions three regions of the valve characteristic, but I think we must in practice make four divisions:—

(i) The temperature-limited region. I have not modified the "pure shot noise" formulae, which I believe are the complete solution of this case.

Table A

Cathode material	Operating temperature	Work-function	kT/e
	°K.	electron-volts	volts
W	2 500	4·52	0·214
W-Th	1 900	2·77	0·163
W-O-Cs	810	0·7	0·069
BaO, SrO, Cao	1 040 to 1 100	0·95	0·089 to 0·094

(ii) The region in which the current is partially but not completely space-charge-limited. This is the practical working condition in amplifying valves, and there is no theory which has been shown to be in agreement with experiment over this region. But the theory given in this paper can be extended to give some idea of the mixed case; this work is in hand and will be published in due course.

(iii) Completely space-charge-limited region. The values of noise calculated by my theory are in close agreement with those predicted by Prof. Schottky. So far as I am aware, this state has never been accurately realized experimentally; but the results of measurements made by Prof. Schottky's colleagues are in moderate agreement with the theoretical limiting value.

(iv) Retarding-field region. For the reasons given above, in reply to Dr. Williams, I do not regard the retarding field region as susceptible to such simple analysis as is sometimes suggested.

I agree that a different effective resistance would have been obtained by considering I^2R in place of E^2/R, but I believe the latter is correct. Any resistance must in practice be shunted by a capacitance, even if only its own self-capacitance; the fluctuation current then charges this capacitance to a corresponding potential, and the capacitance then discharges through all circuit elements which may be connected in parallel across it. The energies dissipated in each are therefore distributed according to factors of the type E^2/R. I have already dealt with the shot-noise interpretation of the retarding-field region, and the reason for describing the component of noise which is proportional to the resistance and temperature of the valve as thermal noise, whatever the mechanism. In saying that the slope resistance requires no modification, it would have been more helpful had Mr. Aldous mentioned which of equations (6) to (10) of the paper he considers to be in error.

In Section (7) of the paper, it will be seen that the mean energy $kT/4$ was found by integrating the energy over all electrons present and dividing by the total number present; it therefore implicitly takes into account the actual law of distribution of velocities, which is admittedly not Maxwellian.

I would willingly endorse the use of conductances rather than resistances, but when comparisons with previous work are involved this might cause difficulty to the reader.

The solution of Dr. Rawlinson's intriguing problem is to be found in the conditions necessary for equipartition to occur. The only application of equipartition in the paper is to the energies of the electrons in the immediate neighbourhood of the cathode, where the free path is likely to be of short duration since the density is very high. Beyond the potential barrier, which is the region where the temperature was stated to be $T/2$, no such equipartition is assumed, and in fact it is improbable that there are sufficient collisions between electrons and molecules, compared with the rate of energy exchange due to collisions between gas molecules and the envelope of the valve, to establish equipartition between the residual gas and the electron stream. In the extreme case of uniform emission, the electron stream would be initially at zero temperature, but would be heated by the residual gas in so far as there were collisions to give a tendency towards equipartition.

I am indebted to Mr. Percival for pointing out the error in equation (21) in the proof copies of the paper; this has now been corrected. In the retarding field region, if $1/b = V$, then $R' = R$ (for evidence that this is not remote from the practical case, see my reply to Dr. Williams on this subject). The argument that electrons near the anode can influence subsequent emission does not destroy the fact that they are not themselves in thermal equilibrium with those near the cathode.

In reply to Prof. Schottky, I have not found it necessary to assume a Maxwellian distribution of velocities in the outer space-charge (see my reply to Mr. Aldous). According to equation (13) of the paper, my theory may be said to give an equivalent temperature of $1\cdot2T/2$ (not $T/2$) for a valve of 3/2 power law, since my expression is $V^2 = 4(1\cdot2R)kT/2$, against Prof. Schottky's $1\cdot272T/2$; I think the residual difference of 6 % could be due to the approximation made at the end of Section (7).

The scope of the paper is to derive the fluctuation noise to be expected from a metallic resistance and from thermionic valves obeying certain ideal laws. The latter cannot in practice be constructed and measured, but this theory of the simpler cases prepares the ground for a treatment of the more complex conditions corresponding to practical valves, which should be both of material value and capable of experimental verification.

A RESONANCE BRIDGE FOR USE AT FREQUENCIES UP TO 10 MEGACYCLES PER SECOND*

By Professor C. L. FORTESCUE, O.B.E., M.A., Member, and G. MOLE, Ph.D., B.Sc.

(*Paper first received* 22*nd September, and in final form* 18*th November*, 1937; *read before the* WIRELESS SECTION 2*nd February*, 1938.)

SUMMARY

This paper describes a form of alternating-current resonance bridge suitable for measuring circuit resistances up to 30 ohms at frequencies up to 0·5 megacycle per sec. and up to 10 ohms at frequencies up to 10 megacycles per sec. The complete theory of the bridge and analysis of the errors is given, together with an account of a series of special check measurements designed to reveal the accuracy actually obtainable. The limit of error appears to be in the neighbourhood of 0·01 ohm for resistances up to 10 ohms at frequencies up to 10 megacycles per sec.

CONTENTS

(1) Principles.
(2) Rearrangement for use at High Frequency.
(3) Constructional Details.
(4) Additional Apparatus.
(5) Wagner Earth System.
(6) Complete Theory of Bridge and Wagner Earth.
(7) Calibration.
(8) Corrections.
(9) High-frequency Operation.
(10) High-frequency Test Circuit.
(11) Results of High-frequency Tests.
(12) Use of Bridge.
(13) Further Development.
(14) Acknowledgments.
(15) Bibliography.

(1) PRINCIPLES

The principle of the resonance bridge, which is well known, is shown in Fig. 1.

Three of the arms of the bridge—R_1, R_2, and R_4—are resistances. The other arm—C, L, R, and R_3—which contains the circuit under test, is tuned to resonance and so behaves also as a pure resistance. The bridge is thus entirely resistive and the conditions of balance are:—

$$LC\omega^2 = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1a)$$

$$R_1R_4 = R_2(R_3 + R) \quad . \quad . \quad . \quad . \quad (1b)$$

(2) REARRANGEMENT FOR USE AT HIGH FREQUENCY

At high frequencies the above conditions are modified by the presence of residual capacitances and inductances, and by the difficulty of constructing satisfactory variable resistances. In the present bridge the effects of residuals have been reduced by making the resistance arms in the form of fine wires mounted axially in screening tubes, and the difficulty with respect to the variable resistance is overcome by the use of a screened straight wire and sliding contact as shown in Figs. 2, 3, and 4. A Wagner earth is used which brings the sliding contact *f* to the potential of the screen. Neither its capacitance to earth nor the resistance of the contact affects the results under these conditions.

The ratio arms R_1 and R_2 are made in the form of a single wire mounted in a screening tube, with a soldered connection to the exact mid-point at *b*.

Still neglecting residuals, the conditions of balance become:—

$$LC\omega^2 = 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2a)$$

$$R = R_4 - R_3 = 2\rho s \quad . \quad . \quad (2b)$$

where ρ is the resistivity of the slide-wire and s is the distance of the contact from the mid-point of the slide-wire. The maximum resistance which can be measured is thus equal to the slide-wire resistance, and is in this case approximately 10 ohms. The range of resistance measurement with this bridge can be extended to 30 ohms at frequencies below 0·5 Mc./sec. by inserting an extra length of wire in the arm *fc*. This operation is not practicable at higher frequencies owing to the terminal capacitances at intermediate points in the arm.

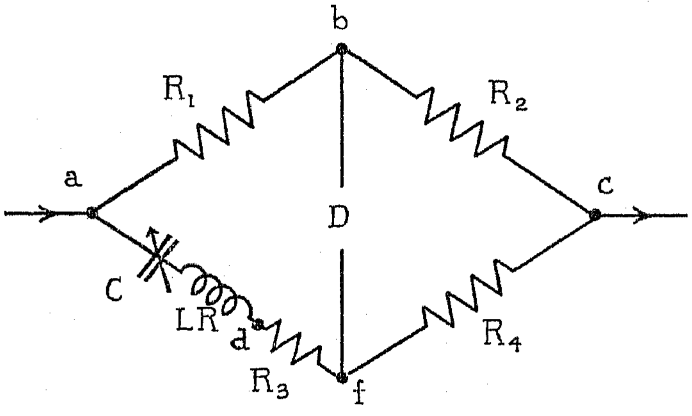

Fig. 1

(3) CONSTRUCTIONAL DETAILS

A complete circuit diagram of the bridge, Wagner earth, and circuit under test, is given in Fig. 2. A diagrammatic plan showing the actual construction is given in Fig. 3.

The apparatus consists of three cylindrical brass tubes 4 cm. diameter and 65–70 cm. long, and a rectangular copper box 80 cm. long and 9 cm. × 6 cm. cross-section. There are also smaller copper boxes at the ends, the whole

---

* Reprinted from *Journal I.E.E.*, 1938, vol. 82, p. 687.

being soldered together to form the screening system of the bridge. The points *a*, *b*, *c*, *d*, *f*, on the bridge are exposed terminals, and the point *e* of Fig. 2 is the screen itself. The terminals are insulated with mycalex and have earth-capacitances of 4·5 $\mu\mu$F each.

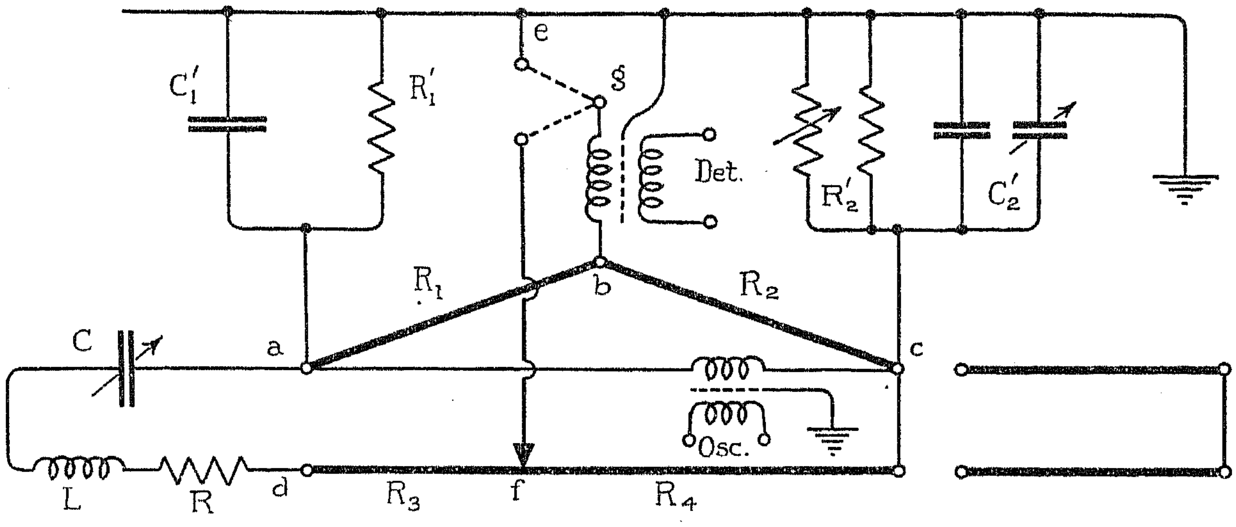

Fig. 2

The ratio arms *ab* and *bc* consist of a single length of No. 36 S.W.G. eureka wire mounted axially in the outer screening tube. The extra resistance used for extending the range of the bridge at low frequencies consists of two lengths of No. 36 S.W.G. eureka wire mounted in the two inner screening tubes. The ends of these wires are connected in the small copper box remote from *c*. The small copper box at the end *c* houses a system of three plug contacts by means of which the extra resistance can be either inserted in, or removed from, the bridge arm *fc*.

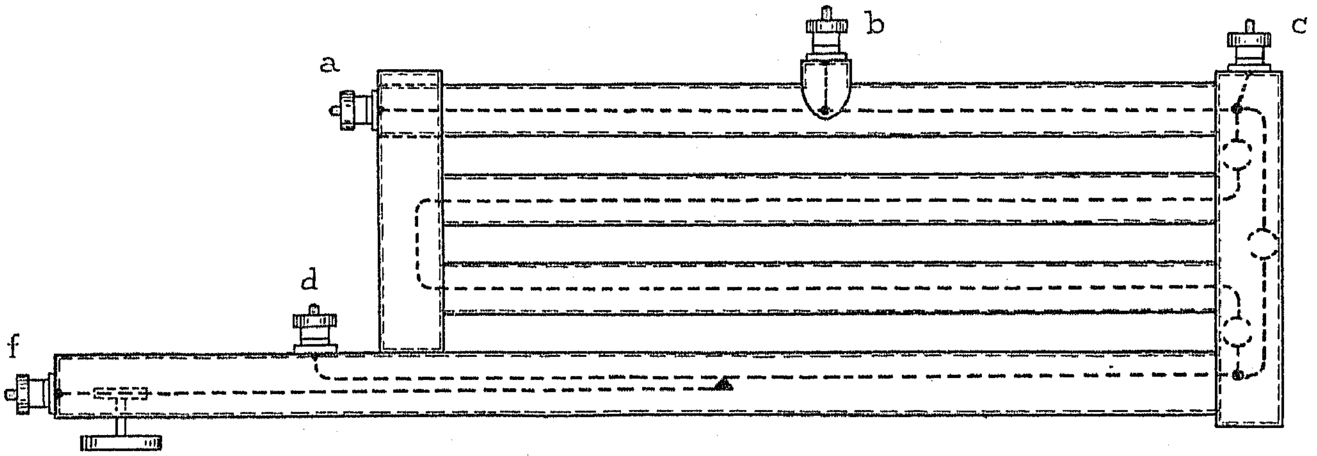

Fig. 3

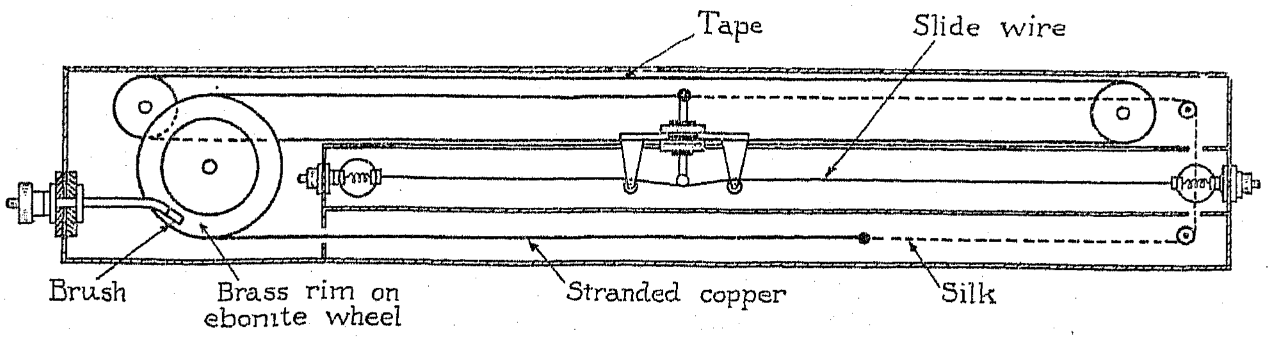

Fig. 4

The long copper box is divided longitudinally into three compartments as shown diagrammatically in Fig. 4. In the middle compartment is the slide-wire, consisting of No. 36 S.W.G. eureka wire equal in length to the ratio arms. In the upper and lower compartments is the variable-length connection between the terminal and the moving contact on the slide-wire *f*. A steel tape is arranged to run around the inside of the upper compartment, closing the longitudinal slot between the upper and middle compartments throughout the whole of its length except for a small insulating bush through which the connection to the contact passes. The steel tape runs over pulleys and is operated by an external handle, graduations on the tape being used to determine the position of the sliding contact. Glass rollers situated on either side of the contact wire and carried with it serve to bend the eureka wire slightly and so maintain the two wires in contact.

The eureka slide-wire is attached to its terminals in the manner shown in Fig. 4. The spring serves to maintain the wire taut under all conditions, and the ring of thin copper strip provides a low-impedance path for the current.

The Wagner earth system consists of fixed and variable condensers and resistances connected in parallel between the terminals *a* and *c* and the nearest points on the screen. The fixed condensers and resistances are arranged to plug in, and are required in values of 100–500 $\mu\mu$F and 1–20 Ω. Of these, the fixed condensers are mica-dielectric condensers of 2 cm. maximum dimensions and possess negligible self-inductance—an essential point since all possibilities of resonance must be avoided. The fixed resistances are carbon-bakelite composition resistances of similar dimensions, also possessing negligible self-inductance. The variable condenser is a small air condenser having a maximum capacitance of 100 $\mu\mu$F and completely screened. The variable resistance is a standard volume control possessing a metallic-film resistance element and having a maximum resistance of 200 Ω.

(4) ADDITIONAL APPARATUS

The additional apparatus required is an oscillator, a detector, and coupling transformers. The construction of the transformers is dependent upon the frequency range in use but is always so arranged that the primary and secondary impedances are approximately matched and the primary and secondary turns are electrostatically screened from each other. The connections between the bridge and the transformers are made with screened leads of equal length and cross-section.

The oscillator is of the ordinary push-pull type and is fed entirely by batteries. The oscillator, batteries, and transformer are enclosed within a copper box and form a unit which is completely screened from the rest of the apparatus. The copper box is itself enclosed within an aluminium screen.

The detector is of the simple heterodyne type and is fed from the mains. It consists of a push-pull, screen-grid, high-frequency amplifying stage; a heterodyne oscillator; a rectifier stage; and a stage of low-frequency amplification. The whole is enclosed within a partitioned aluminium screen.

It is important that there should be no direct coupling between the oscillator and the detector. The elimination of all such coupling is easily proved by disconnecting the bridge from the oscillator and detector leads, when no signal from the oscillator should be receivable. With the arrangement of multiple screening described above, no direct coupling is, in fact, detectable at any frequency up to 10 Mc./sec.

(5) WAGNER EARTH SYSTEM

The bridge is not connected directly to earth at any point, but when it is balanced the points *b* and *f* are exactly at earth potential. Earth capacitances at these points can therefore be neglected. This is essential, since the earth capacitance at *f* depends upon the adjustment of the slide-wire contact. The earth capacitances at *a* and *c* are in parallel with the Wagner-earth condensers C_1' and C_2', and form no part of the bridge circuit. The Wagner-earth system thus eliminates the effects of the earth capacitances at the points *a*, *b*, *c*, *f*.

The only earth capacitances remaining are the distributed capacitances of the wires and the capacitance at *d*. The maximum frequency at which the bridge is intended to be operated is 10 Mc./sec., corresponding to a minimum wavelength of 30 metres. The minimum half-wavelength, 15 m., is thus great in comparison with the length of either the ratio arms or the slide-wire, and consequently the distributed capacitances of the wires can be regarded as being concentrated at the ends, where they are merely parts of the earth capacitances at the points *a*, *b*, *c*, *f*, and *d*.

The effect of the total earth capacitance at the point *d* cannot be eliminated from the bridge. Therefore, either it must be made very small or a correction must be applied for it.

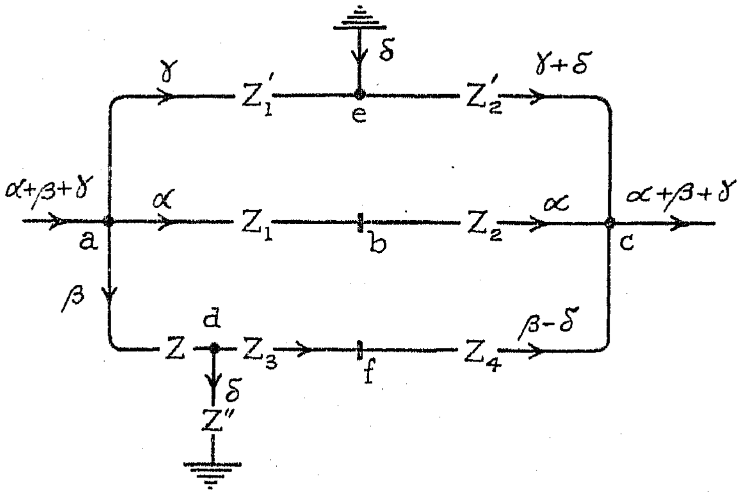

Fig. 5

(6) COMPLETE THEORY OF BRIDGE AND WAGNER EARTH

Fig. 5 is a complete circuit diagram of the bridge and Wagner earth, including the earth capacitance at the point *d*. The symbols Z_1, Z_2, Z_3, Z_4, Z_1', Z_2', Z'', Z represent the complex values of the impedances constituting the network, whilst $\alpha, \beta, \gamma, \delta$ represent the complex values of the currents in the network.

A balance occurs when the points *e*, *b*, and *f* are at the same potential. Since the two ratio arms are made equal, $Z_2 = Z_1$. Therefore,

$$Z_1\alpha = Z_1'\gamma \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3a)$$

$$Z_1\alpha = Z\beta + Z_3(\beta - \delta) \quad . \quad . \quad . \quad (3b)$$

$$Z_1\alpha = Z\beta + Z''\delta \quad . \quad . \quad . \quad . \quad . \quad (3c)$$

$$Z_1\alpha = Z_2'(\gamma + \delta) \quad . \quad . \quad . \quad . \quad (3d)$$

$$Z_1\alpha = Z_4(\beta - \delta) \quad . \quad . \quad . \quad . \quad . \quad (3e)$$

Eliminating Z_1, α, β, γ, and δ,

$$\frac{1}{Z_2'} - \frac{1}{Z_1'} = \frac{Z_3}{Z_4 Z''} \quad . \quad . \quad . \quad . \quad . \quad (4a)$$

$$Z = \frac{Z_4 - Z_3}{1 + \dfrac{Z_3}{Z''}} \quad . \quad . \quad . \quad . \quad (4b)$$

When separated into real and imaginary parts, equations (4*a*) and (4*b*) become:—

$$\frac{1}{R_2'}-\frac{1}{R_1'}=\frac{(R_3L_4-R_4L_3)C''\omega^2}{R_4^2+L_4^2\omega^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5a)$$

$$C_2'-C_1'=\frac{(R_3R_4-L_3L_4\omega^2)C''}{R_4^2+L_4^2\omega^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5b)$$

$$R=\frac{R_4-R_3+(R_4L_3-R_3L_4)C''\omega^2}{(1-L_3C''\omega^2)^2+R_3^2C''^2\omega^2} \quad . \quad . \quad . \quad . \qquad (6a)$$

$$L\omega-\frac{1}{C\omega}=\frac{(L_4\omega-L_3\omega)(1-L_3C''\omega^2)-(R_4-R_3)R_3C''\omega}{(1-L_3C''\omega^2)^2+R_3^2C''^2\omega^2} \qquad (6b)$$

It will be seen that equations (5*a*) and (5*b*) represent the conditions of balance for the Wagner earth, whilst equations (6*a*) and (6*b*) represent the conditions of balance for the bridge. The equations (5*a*) to (6*b*) are greatly simplified if it is permissible to assume that the inductance and resistance of the end connections are relatively small compared with those of the wires and that, in addition, the ratio of the resistance to the inductance of the end connections does not differ greatly from that of the wires. Under these conditions,

$$R_3L_4 = R_4L_3$$

and

$$\frac{(R_4-R_3)R_3C''}{L_4-L_3}=\frac{R_3^2C''}{L_3}$$

Equations (5*a*), (5*b*), (6*a*), and (6*b*) therefore reduce to:—

$$R_1' = R_2' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (7a)$$

$$C_2'-C_1'=\frac{(R_3R_4-L_3L_4\omega^2)C''}{R_4^2+L_4^2\omega^2} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (7b)$$

$$R=\frac{R_4-R_3}{(1-L_3C''\omega^2)^2+R_3^2C''^2\omega^2} \quad . \quad . \quad . \quad . \qquad (8a)$$

$$L\omega-\frac{1}{C\omega}=\frac{(L_4\omega-L_3\omega)\left(1-L_3C''\omega^2-\dfrac{R_3^2C''}{L_3}\right)}{(1-L_3C''\omega^2)^2+R_3^2C''^2\omega^2}. \qquad (8b)$$

In the case of the present bridge with a value of C'' of 40 $\mu\mu$F, the maximum values of the terms $R_3^2C''^2\omega^2$ and R_3^2C''/L_3 are respectively 0·00015 and 0·0003, and may be neglected in comparison with unity. Equations (8*a*) and (8*b*) thus become:—

$$R=\frac{R_4-R_3}{(1-L_3C''\omega^2)^2} \quad . \quad . \quad . \quad . \qquad (9a)$$

$$L\omega-\frac{1}{C\omega}=\frac{(L_4-L_3)\omega}{1-L_3C''\omega^2} \quad . \quad . \quad . \quad . \qquad (9b)$$

The bridge balance is represented by equations (9*a*) and (9*b*) with a maximum error of 1 part in 2 000. Inspection of equations (7*a*) and (7*b*) shows that the adjustment of the Wagner earth is dependent upon the slide-wire adjustment.

(7) CALIBRATION

The bridge was calibrated with direct current by connecting a standard variable resistance between the terminals *a* and *d*.

(8) CORRECTIONS

For resistance measurements having an accuracy of the order of 2 %, it is possible to use directly the d.c. calibration of the bridge. When more accurate measurements are to be made, it is necessary to apply a correction for skin effect in the slide-wire. The values of the correction in the case of the present bridge are given in Table 1, where $R_{D.C.}$ is the test-circuit resistance as determined from the bridge balance and the d.c. calibration and R_F is the test-circuit resistance after correction for skin effect in the slide-wire.

Table 1

f (in Mc./sec.) =	0·70	1·30	2·00	3·00	5·00	7·50	10·00
$R_F/R_{D.C.}$ =	$1\cdot000_1$	$1\cdot000_2$	$1\cdot000_4$	$1\cdot001_0$	$1\cdot002_9$	$1\cdot007_3$	$1\cdot013_2$

Where the effect of the capacitance C'' is appreciable, it is necessary to use the relations (9*a*) and (9*b*). These may be rewritten:—

$$R_E = R = \frac{R_F}{(1-L_3C''\omega^2)^2} \quad . \quad . \qquad (10a)$$

$$L\omega-\frac{1}{C\omega}=\frac{\omega\mu R_{D.C.}}{1-L_3C''\omega^2} \quad . \quad . \quad . \qquad (10b)$$

where μ represents the ratio of inductance to d.c. resistance in the slide-wire. For the present bridge the value of μ is 0·068 μH per ohm, and the value of L_3 is given (in μH) by

$$L_3 = 0\cdot354 - 0\cdot034R_{D.C.}$$

(9) HIGH-FREQUENCY OPERATION

The circuit under test should be connected across the terminals *a* and *d* on the bridge in such a way that the part of the circuit possessing the greatest earth capacitance is connected to the point *a*. The earth capacitance at *d* should be as low as possible.

The bridge and Wagner earth require four adjustments for balancing. They are (i) the adjustment of the Wagner-earth variable resistance R_2', (ii) the adjustment of the Wagner-earth variable capacitance C_2', (iii) the adjustment of the slide-wire contact, and (iv) the adjustment of the variable reactance in the test circuit. Of these, (i) is independent, but (ii), (iii), and (iv) are interdependent. Adjustment (ii) has, however, only a small effect on (iii) and (iv). The final adjustment is indicated by a minimum of sound in the telephone, but if the adjustment of the slide-wire contact is not near its correct value then the reactance adjustment (iv) should be to a maximum instead of a minimum, since it is then a form of resonance adjustment. It is possible to obtain a perfect balance of both bridge and Wagner earth over the whole range of frequency for which they are intended to be used, i.e. from 0 to 10 Mc./sec.

(10) HIGH-FREQUENCY TEST CIRCUIT

A circuit employed for testing the accuracy of the bridge over the frequency range 1–10 Mc./sec. is shown in Fig. 6. It consisted of a standard, variable, air condenser, a doubly-screened coil, and a doubly-screened resistance unit. The air condenser was one of two, possessing maximum capacitances of 1 200 $\mu\mu$F and 200 $\mu\mu$F and having very low power factors. The coils were available in values ranging from $0\cdot5$ μH to 50 μH.

The resistance took the form of a three-terminal, screened, plug-in holder arranged to take one of two resistance units. These two units consisted of equal $15\cdot5$-cm. lengths of No. 40 S.W.G. copper and eureka wires of uniform and equal diameters, mounted in similar brass screening tubes. The two units thus possessed equal self-inductance and self-capacitance but differed in resistance. Substitution of one unit for the other, therefore, would produce a resistance-change in the test circuit but would not in itself change the reactance of the circuit. The readjustment of the slide-wire contact necessitated by the resistance-change would, however, require to be balanced by a capacitance-change in the test condenser. In order to ensure that this capacitance-change would not have associated with it an extra, small, resistance-change, it was necessary that the power factors of both the condenser and the capacitance between the inner and outer coil screens should be low.

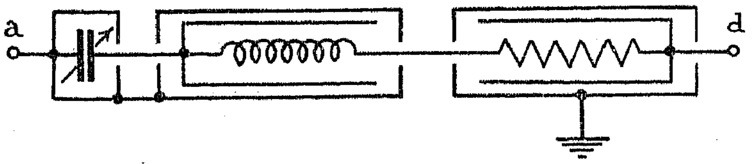

Fig. 6

The direct-current resistances of the units, $R_{d.c.}$, were measured with the bridge itself, using a current of only 5 mA in order to avoid thermal resistance-change in the copper unit. A current of the same order of magnitude was employed in the high-frequency tests.

In order to determine the effective resistances, R_e, of the units it was necessary to apply corrections for (i) skin effect (giving R_f) and (ii) the effect of the self-inductance l and self-capacitance c of the units. The latter correction was given by

$$R_e = \frac{R_f}{(1 - lc\omega^2)^2}$$

Other residuals in the circuit were treated as follows: The capacitance between the inner and outer resistance-unit screens merely added to the existing earth capacitance C'' at d and could be taken into account by using the exact bridge-balance relations (10a) and (10b). The connection between the coil and the resistance unit was made so short that any residual associated with it could be neglected. The capacitance between the inner and outer coil screens merely added to the condenser capacitance. The earth capacitances of the outer coil screen and of the condenser screen formed part of the Wagner earth circuit and so did not affect the bridge.

(11) RESULTS OF HIGH-FREQUENCY TESTS

The bridge was tested over a frequency range of 1–10 Mc./sec. using the test circuit described above. The effective high-frequency resistance of the test circuit was measured with the copper resistance unit and again with the eureka resistance unit. The difference gave the difference between the effective high-frequency resistances of the two units and could be compared with the known value of this difference as determined by direct-current measurement and correction. The results of the tests are given in Table 2.

Table 2

Frequency	Test coil	Resistance difference (Eureka — Cu)			Percentage error
		Test resistance value	Bridge value	Error	
Mc./sec.		Ω	Ω	Ω	
$1\cdot30$	A	$6\cdot30_7$	$6\cdot31_9$	$+\ 0\cdot01_2$	$+\ 0\cdot2$
	a	$6\cdot30_7$	$6\cdot27_2$	$-\ 0\cdot03_5$	$-\ 0\cdot5$
	a_2	$6\cdot30_7$	$6\cdot29_8$	$-\ 0\cdot00_9$	$-\ 0\cdot1$
$2\cdot00$	a	$6\cdot30_3$	$6\cdot30_2$	$-\ 0\cdot00_1$	$-\ 0\cdot0$
	a_2	$6\cdot30_3$	$6\cdot28_0$	$-\ 0\cdot02_3$	$-\ 0\cdot4$
$3\cdot00$	a_2	$6\cdot29_3$	$6\cdot28_7$	$-\ 0\cdot00_6$	$-\ 0\cdot1$
	7t	$6\cdot29_3$	$6\cdot28_8$	$-\ 0\cdot00_5$	$-\ 0\cdot1$
$5\cdot00$	7t	$6\cdot27_1$	$6\cdot23_9$	$-\ 0\cdot03_2$	$-\ 0\cdot5$
	4t	$6\cdot27_1$	$6\cdot27_9$	$+\ 0\cdot00_8$	$+\ 0\cdot1$
$7\cdot50$	7t	$6\cdot25_3$	$6\cdot24_9$	$-\ 0\cdot00_4$	$-\ 0\cdot1$
	4t	$6\cdot25_3$	$6\cdot27_6$	$+\ 0\cdot02_3$	$+\ 0\cdot4$
	2t	$6\cdot25_3$	$6\cdot27_3$	$+\ 0\cdot02_0$	$+\ 0\cdot3$
$10\cdot00$	4t	$6\cdot25_5$	$6\cdot22_7$	$-\ 0\cdot02_8$	$-\ 0\cdot4$
	2t	$6\cdot25_5$	$6\cdot27_6$	$+\ 0\cdot02_1$	$+\ 0\cdot3$

It will be seen that the mean error in the measurement of a resistance of $6\cdot3\Omega$ is $0\cdot01_6\Omega$, or $0\cdot25$ %. Since this represents the sum of the errors in two measurements, it can be said that the bridge is capable of measuring resistances of 0–10Ω over the frequency range 1–10 Mc./sec. with a probable error of $\pm\ 0\cdot01\Omega$. It appears from Table 2 that the errors are randomly distributed over the frequency range. The accuracy of the bridge does not, therefore, appear to diminish with increasing frequency.

(12) USE OF BRIDGE

The bridge can be used for obtaining accurate comparisons of the resistance and reactance of any circuit at frequencies up to 10 Mc./sec., providing (i) the equivalent series resistance does not exceed 10Ω (30Ω for frequencies below 0·5 Mc./sec.), (ii) the reactance is finely adjustable to resonance, and (iii) the circuit does not possess excessive capacitance to earth at points other than its terminals. Where there is choice in the arrangement of screens it is possible completely to eliminate earth capacitance of internal parts of the test circuit, as in the circuit of Fig. 6.

A useful test circuit for measurement of the reactance and resistance of screened coils is shown in Fig. 7. The capacitance at the point *d* can be made sufficiently small to render correction for it unnecessary. If a low-power-factor condenser is used, the equivalent series resistance of the coil is obtained directly from the slide-wire calibration with a limit of error of 1 % and, after correction for skin effect only, with greater accuracy. The reactance is given by

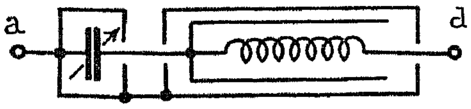

Fig. 7

$$\frac{1}{C\omega} + \omega\mu R_{D.C.}$$

where *C* is the condenser capacitance. It should be observed that only when one terminal of the coil is connected to the screen are the effective values of reactance and resistance independent of the circuit in which the coil is connected.

The bridge is, in general, found to be more sensitive than the existing resonance methods for measurement of reactance and resistance over the 1–10 Mc./sec. frequency range.

(13) FURTHER DEVELOPMENT

There is at present under construction a bridge similar in principle to the above but of smaller dimensions, which is intended to cover the frequency range 10–100 Mc./sec.

(14) ACKNOWLEDGMENTS

Much of the preliminary work on this bridge was carried out by postgraduate students of the City and Guilds College, South Kensington, notably Dr. Z. U. Ahmad and Mr. R. P. Glover. This work prepared the way for later developments. The completion of the work was rendered possible by the generous grants for research made to the College by the Worshipful Company of Clothworkers.

(15) BIBLIOGRAPHY

Details of previous forms of radio-frequency bridge will be found in the following papers:—

(1) T. I. Jones: "A Critical Résumé upon High-Frequency Bridge Networks," E.R.A. Report Ref. L/T56, 1933.
(2) D. W. Dye and T. I. Jones: "A Radio-Frequency Bridge for Impedance and Power-Factor Measurements," *Journal I.E.E.*, 1933, vol. 72, p. 169.
(3) W. J. Shackleton: "A Shielded Bridge for Inductive Impedance Measurements," *Bell System Technical Journal*, 1927, vol. 6, p. 142.
(4) J. G. Ferguson: "Shielding in High-Frequency Measurements," *ibid.*, 1929, vol. 8, p. 560.

DISCUSSION BEFORE THE WIRELESS SECTION, 2ND FEBRUARY, 1938

Dr. L. Hartshorn: The sensitivity and perfection of balance of the bridge can, in these days of amplifiers, and efficient screening and earthing, be taken for granted. The practical value of the new bridge will therefore depend on its overall accuracy, convenience in use, and flexibility, i.e. the ease with which it may be adapted to deal with jobs of various kinds. It would be of great value if the authors could give some idea as to how the bridge compares with alternative methods in these respects. The estimated limit of error of 1 in 1 000 is probably also the upper limit of accuracy obtainable by the resonance method. The fact that the reactance is tuned for a maximum of sound in the initial stages, and a minimum in the final stages, suggests that the adjustments might at times be less straightforward than those of untuned bridges, but it may be that the apparent ambiguity causes no practical inconvenience.

The resonance bridge was used some years ago at the N.P.L. at audio frequencies, but was discontinued for measurements on ordinary coils, owing to the fact that at resonance the junction of coil and condensers is at very high potential, and that therefore errors due to any earth-capacitance currents at this point are apt to be unduly large. The Wagner earth only allows for capacitance currents at the corners of the bridge; the bridge readings are therefore a function of the earth capacitance at the junction in question, and are in consequence limited in their application to the special conditions imposed by the bridge network. The authors' scheme of connections shown in Fig. 6 is exactly that required to eliminate error due to this cause: the screens are so arranged that the junction has capacitance to the condenser screen-terminal but not to earth. I regard this feature as essential for accurate results. Does not the necessity for such conditions limit the application of the method somewhat severely? Measurements on unscreened coils would, I think, from my own experience, be subject to larger errors if made by this method. Nevertheless, it is of the utmost importance that the full possibilities of such methods should be thoroughly explored, and in showing that the method is capable of an accuracy within 0·1 % at 10 Mc./sec. under suitable conditions the authors' investigation is a valuable piece of work.

Mr. T. Iorwerth Jones: It is interesting to recall that the earliest form of radio-frequency bridge, described in this country about 15 years ago by Mr. M. Hart, was also developed at the City and Guilds Engineering College. In some respects it resembled the bridge described in the present paper. Its arms consisted similarly of eureka wires, rather longer, of course, than those in the authors' apparatus. A movable contact was realized by threading the wire through a small cube containing mercury, the holes being so small that the mercury did not escape. There was no screening. The full implications of the pioneer work of G. A. Campbell, and its after development at the hands of Shackleton and Ferguson, of the Bell Telephone Laboratories, were realized subsequently. When we came to design a radio-frequency bridge at the N.P.L., in adapting a bridge to radio frequencies we introduced a great deal of screening; we were compelled to do so, because we were aiming at the measurement in the first place of very, very small phase angles associated with the effective resistances of condensers, phase angles which can frequently be as low as 6 sec. of arc at 1 Mc./sec. The bridge was also intended for measurements on dielectrics. In those measurements the apparatus was provided with two electrodes and a guard-ring, and the potential of one of those electrodes had to be maintained accurately at the potential of the guard-ring, otherwise a different result would be obtained, depending upon the choice of the arm into which the apparatus was inserted. Screening and the elimination of induction into the detector circuit were the essential precautions which had to be observed.

The authors' bridge is essentially a bridge with low-impedance arms. It is assumed that it is not intended for the measurement of very small phase-angles. This allows of greater latitude in the matter of screening—at the upper reaches of the frequency range this becomes imperative if the bridge is to remain workable, otherwise the earth admittances would exceed the measured admittances and complicate the result. There is a definite limit to the accumulation of earth capacitances above 1 Mc./sec. The authors' bridge appears, in general, to be admirable for the measurement of screened inductances and for the adjustment to resonance at a definite frequency of screened composite units intended for filter circuits, and as a means of deriving a knowledge of the resistance of such components. There is one feature, however, on which one is not too happy. The tapping on the variable resistance carries a detector point which moves a considerable distance along the wire between the two balances obtained with the apparatus under test included in, and removed from, the bridge arm respectively. This movement may modify the mutual induction between the source circuit and the detector circuit which may almost inevitably exist to a small extent. It is difficult to estimate this effect from a cursory glance at the assemblage.

A few years ago, in the course of a review* of high-frequency bridge networks, I had occasion to comment upon the absence of satisfactory standards of resistance for work at high frequency. The authors' bridge embodies standard resistances both of the fixed and of the variable types which are admirably screened and compact.

Mr. E. B. Moullin: The authors use a battery-driven generator and mains-operated receiver, or vice versa. I presume they fear to use two mains-driven pieces of apparatus lest a signal should pass from one to the other through the a.c. system. I should like to know whether they think it is beyond their ingenuity to arrange for both units to have a.c. drive.

I notice that the authors use mycalex for terminal bushings; had they any reason for using it in preference to, say, trolitol, which I myself prefer?

It is a little difficult to assess the order of importance of C'' from the two bridge-balance conditions, equations (10*a*) and (10*b*). Its action is that the equivalent of inserted resistance in Fig. 5 is shunted by a capacitance, so that its effective resistance is smaller than it ought to be; the difference is enhanced by the residual inductance of the arm. I feel that the effect of C'' is probably very small indeed, because L_3 can be made very small. Why was no test made by connecting a variable capacitance to earth from the danger point *d*, and examining the inaccuracy produced?

It is stated in Section (10) that the inductances of units of copper wire and eureka wire are the same. I maintain there is a certain range of frequency in which the two values will not be the same, because the inductance due to the internal field is a function of frequency and is not a negligible fraction of the total inductance. I suggest that the inductance for steady currents would be about 10 % more than the limiting final value. The rates of approach to that lower value would be different for the copper and the eureka, having two different penetrations. As I presume that the term L_3 in equation (10*a*) is negligible anyhow, this discrepancy does not much matter.

It would be valuable if other speakers in the discussion would give their views as to the relative merits of the use of a bridge and of straightforward resonance methods. I hope it is still reasonable to go on using resonance rather than bridge methods, and I gather that Dr. Hartshorn thinks likewise. What do we gain by using a bridge method? Essentially that we cut out any indicating instrument the law of whose scale we have to rely on. We throw all the responsibility on to a standard resistance or condenser: provided these remain constant then obviously the bridge method gives a much higher degree of possible accuracy of reading than the resonance method. My experience of resonance methods has not revealed a lack of accuracy of reading, desirable as it may be to have the further extended accuracy of reading which the authors' bridge must undoubtedly give. There is, however, one point which is worth bearing in mind. So far as my experience goes, the high-frequency resistance of a circuit is not constant from day to day to the extent of 2 parts in 1 000. It is easy to make a coil which has a power factor of 0·3 %, and the operation becomes successively easier as one goes to higher and higher frequencies. That being so, the power factor of the condenser becomes an appreciable fraction, of the order of 10 %, of the whole. I think that in ordinary circumstances the power factor of condensers, with their bushings affected by conditions of weather, is slightly variable from day to day. It is certainly my opinion that there is not a greater day-to-day constancy than about 0·5 %, so that I think there is an upper limit to which it is useful to go for precision of comparison.

A further point which occurs to me is that the authors' bridge essentially measures resistance, whereas in general the quantity which is more important in high-frequency work is the power factor of the circuit.

Mr. Charles Holt Smith: I have throughout the last year been developing a bridge which is complementary to the authors'. The conditions which I had to fulfil were that it should be suitable for use in the field at Daventry for the purpose of measuring and adjusting short-wave transmitter aerials. It had therefore to be easily portable, and weatherproof. It had to operate over frequencies between 6 Mc./sec. and 24 Mc./sec. and to be suitable for measuring balanced impedances. Fortunately, a limit of error of the order of 5 % was considered ample. This bridge was duly constructed, and it served a very useful purpose throughout the past year in connection with the setting-up of the Daventry aerials.

---

* E.R.A. Report L/T 56, 1933.

The difficulties experienced by residual capacitance at the junction point of a standard resistance and condenser were avoided by connecting the standard resistance and condenser in parallel, so that the answer obtained for the unknown impedance was expressed as the equivalent parallel resistance and reactance. The reactance was measured by means of a double differential condenser. Differential condensers were employed in order that by transferring negative reactance into either the test arm or the standard arm of the bridge, positive or negative reactance could be measured, and two condensers were used in series to give a balanced system with the centre-point earthed. The resistance component of the impedance to be measured was determined on a carbon track varying with the resistance. The resistance used was a hand-calibrated Morgan–Stackpole 1 000-ohm resistance. The resistance track was of a form conforming to no previously known rule, having been designed in such a manner that the scale was open mainly in the region round about 600 ohms, at which value maximum accuracy was required. A further device which was incorporated permitted the whole of the range from zero to infinity shunt resistance to be included on a single dial. By connecting across the test terminals of the bridge a variable resistance, and by means of an initial bridge adjustment whereby this resistance was made equal to the maximum value of the variable resistance in the standard arm of the bridge, the external shunt resistance across the test arm under conditions of balance was given the value infinity; whereas when the standard resistance was turned down to zero the total resistance variation was from zero to infinity. The reactance variation was of the order of ± 70 $\mu\mu$F.

Dr. M. Reed: Have the authors been able to verify that the magnitude of the contact resistance which would be at f if the whole of the variable resistance was confined to R_4 is sufficient to justify the considerable complication caused by leaving d with an impedance to earth?

Although the use of cylindrical tubes does seem to give an ideal layout from the electrical point of view, it is interesting to consider whether equal results could not have been achieved by a more compact form of bridge. For example, have the authors investigated the possibility of (*a*) using for their ratio-arms balanced coils wound with twisted pair on a core of ferromagnetic material suitable for operation at high frequencies, (*b*) having a spool consisting of an Ayrton–Perry winding of eureka wire on a small card of, say, 1 in. square for the additional fixed resistances?

In connection with the formula for the reactance correction given in (10*b*), have the authors considered the possible inductance of the air condenser, and, if so, is its magnitude too small to be included in the correction term? Finally, since the errors given in Table 2 seem to be independent of frequency, what decides the limiting frequency at which the bridge can be operated?

Mr. N. Lea: I should be interested to learn whether the authors have made any assessment of the errors they would expect beyond the limit which they have imposed on the working frequency of their bridge, namely 10 Mc./sec.

A further point is that the figure of 4·5 $\mu\mu$F mentioned for the capacitance of terminals is a little higher than one would expect to be necessary. Have the authors tried to reduce this value?

I am not quite sure whether Mr. Moullin's remarks about an alternative method of measurement referred to the resonance method wherein one injects a known voltage into the circuit under test and then measures the voltage across the circuit. If so, I agree with him that it is possible to get all the discrimination which one needs with that arrangement, the limit of error being less than 1 in 1 000. Some time ago, I had occasion to carry out some radio-frequency loss tests on good-quality insulating oils, and I found that with quite reasonable precautions one could detect changes of the order of 0·0001 ohm in about 5 ohms.

Dr. Hartshorn has made a very strong point about the danger associated with the mid-point between the condenser and the inductance in the circuit to be measured. This is of even more importance in commercial work, where often we have to deal with quite large coils and condensers.

Dr. A. Rosen: The authors state that their bridge is to be regarded as suitable for reactance as well as resistance measurements; the variable condenser C is consequently an essential part, and more detailed information concerning it might be included in the paper, possibly in Section (4). In that case, the condition that the reactance of the measured circuit should be finely adjustable could be omitted from Section (12). On the other hand, I would suggest that the following limitations be added: (*a*) The measured circuit must not be earthed at any point. (*b*) It should be completely shielded, as shown in Fig. 7, preferably with an additional earthed screen to make the capacitance to earth constant. (*c*) The capacitance between the screens, which is thrown across the variable condenser C, should be low enough to permit the circuit being brought to resonance. (*d*) The dielectric between these screens should be of the highest quality, so that this capacitance has negligible series resistance.

Mr. H. Page (*communicated*): In making measurements on coils by means of an impedance bridge, it is desirable to connect the coil in the same way as that in which it is used in practice. In the case of a series-tuned circuit, for instance, the value of resistance in which we are interested is that obtained by measuring the series resonant impedance. In the case of a parallel-tuned circuit, however, the effective resistance of the coil can only be obtained by measuring the parallel resonant impedance; in this case the self-capacitance of the coil forms part of the tuning condenser and does not affect the resistance measurement. The effect of self-capacitance can, of course, be calculated, but the calculation is laborious and generally inaccurate.

The application of the series resonance bridge is limited, therefore, to coils of small inductance, in which the self-capacitance does not affect the resistance measurement. This excludes the majority of coils used in modern broadcast receivers.

The use of wire mounted axially in screening tubes for the ratio arms is interesting; in the types of bridge normally used for impedance measurements, these take the form of lumped elements, either the inductive type or Ayrton–Perry resistance windings. Two similar units give good equivalence of impedance up to 5 Mc., and possibly higher, but I am not aware of any tests which have been made to the same accuracy and over the same frequency-range as those of the authors. Their experience in this respect would be of interest.

Prof. C. L. Fortescue and **Dr. G. Mole** (*in reply*): Several speakers have raised the question of the nature of the measurements for which this bridge is designed. It was originally intended for the measurement of the equivalent series resistance of any series L—C—R circuit or of any combination of circuits that could be represented by L, C, and R in series at the particular frequency in use. It must be possible to adjust the equivalent reactance, and under the conditions of measurement no part of the circuit under test is actually earthed. Both ends of the circuit under test are, however, very nearly at earth potential, and any leakage from the high-potential point of the circuit will only be that occurring under normal operating conditions. Thus, contrary to the opinions expressed by one or two speakers, no error arises from this cause. It is only when attempting to prove the accuracy of the bridge, as described in Section (10) of the paper, that this leakage has to be reduced to a minimum or by some means maintained constant.

As a by-product, the bridge proved extremely convenient for the comparison of reactive components by the method of simple substitution.

The bridge is not adapted for the measurement of parallel L, C, R circuits, owing to the difficulty of obtaining high resistances of accurately known values and to the intention of keeping all the impedances of the bridge quite low. Thus inductances tuned wholly or nearly wholly by self-capacitance cannot be compared on this bridge. The simple form of bridge described by Mr. Smith illustrates this difficulty, for he claims a limit of error of only ± 5 %, whereas the bridge described in the paper has a proved limit of error approaching 0·1 %.

The screening of the test circuit has been mentioned. Screening is not necessary unless it is the normal operating condition. But if it is not provided the resistance and reactance will depend upon the immediate surroundings, and any re-arrangement will lead to a change of resistance which actually occurs and which is duly recorded by the bridge measurement.

The result given by the bridge is an effective resistance. The frequency is known, and if either L or C is known the power factor of the circuit as a whole can be found. The power factor of either the capacitance or the inductance alone cannot be found, but change of power factor due to change of components is easily observed.

The relative advantages of the bridge measurement and the resonance measurement is naturally the subject of several questions. Mr. Moullin really answers these questions, except that he does not mention the important fact that when the resonance method is used there will either be some uncertainty with regard to the losses or the additional reactance involved in the indicating instrument; or the complication of the Mallett–Blumlein method will be involved. The bridge gives the effective series resistance without these difficulties. It does not give a power factor, but in practice the effective resistance is usually the more important. The bridge measurement is probably slightly more accurate than the resonance measurement, the circumstances being similar.

Mr. Moullin mentions several other important points. It is quite possible that both receiver and oscillator could be operated from the mains, but this plan was not adopted, as it seemed unwise to run any risks in the course of the development of the bridge. The change of inductance of the copper and eureka wires is an important point where this substitution is employed. In the present instance the possible change of L is about $0\cdot008\ \mu$H in a total inductance of the resistance wire of $0\cdot18\ \mu$H, which is, in itself, only a part of the whole inductance of the circuit. The error arising from this cause is thus very small. With regard to the effects of C'', the tests of the bridge given in Section (10) involved capacitances much higher than those to be expected under normal conditions of working. Thus, if the corrections were satisfactory in those tests there is no reason to suppose that greater errors will arise when the capacitance C'' is much less. The material used for the terminal insulators is unimportant, owing to the low impedance of the bridge. Mr. Lea suggests that the terminal capacitance should be less than $4\cdot5\ \mu\mu$F, and he may be interested to know that in the new bridge, designed for higher frequencies, this capacitance is reduced to little over $1\cdot7\ \mu\mu$F.

The limitation of frequency with the original bridge arises from a variety of causes such as the high terminal capacitance mentioned by Mr. Lea, and the distance apart of the terminals to which the test circuit is connected. Primarily, however, the limitation arises from the fact that the series resistances of higher-frequency circuits are too low to be measured accurately on a 10-ohm slide wire of under 1 metre in length.

Two speakers have suggested the use of wound coils for the ratio arms. This is possible, but it is not so easy to ensure their equality under all conditions as in the case of the straight wire tapped at the centre, where a single resistance comparison ensures equality of reactance within the limits of constancy of the wire.

Mr. Jones asks whether variation of the position of the slide-wire contact causes a variation of the coupling with the receiver. The sliding contact is always at earth potential when the bridge is balanced, and it thus seems unlikely that any error can arise from this cause.

DIRECTIONAL RECORDING OF RADIO ATMOSPHERICS*

By F. E. LUTKIN.

[From the National Physical Laboratory.]

(*Paper first received 22nd September, and in final form 22nd November,* 1937.)

SUMMARY

The paper describes an improved recorder for indicating the direction of arrival of transient impulses continuously throughout the day. The polar diagram of the receiver has a width of 20° and is unambiguous. The analysis of three years' records shows that most of the atmospheric disturbance reaching a station in England arrives from the west or south-west, but that during summer afternoons sources to the east and south-east are productive together with storms of a purely local character. The disturbed azimuths appear closely related to the bearing of the great tropical thunderstorm centres, but the intensity of disturbance from each is dependent on the ionization conditions prevailing over the path to the receiver, at the time each region is productive. Curves are given from which the probable intensity of atmospheric in any direction at any time may be roughly determined.

INTRODUCTION

In a paper entitled "Directional Recording of Atmospherics," Watson Watt† has described the recorder used for some time in the work of the Radio Research Board. Its purpose was to register throughout 24 hours the azimuth of arrival of the electrical impulses called variously "atmospherics," "strays," etc. In the course of analysis of the records from this recorder, it became evident that in many cases atmospherics were arriving from more than one direction, and the physical significance of the results was being masked by the inability to determine more than the mean azimuth of disturbance when two or more streams were simultaneously active.

The improvements outlined in the present paper were originally suggested in the summer of 1927, but it was not until 1930 that time was found in the programme to proceed with the development of the instrument. Since then, and until the middle of 1935, the apparatus has been in almost continuous use, and has proved quite stable in operation and satisfactory in result.

It consists essentially of two frame aerials, referred to as the "recording frame" (R) and the "opposing frame" (O), mounted mutually at right angles on a common vertical spindle (see Fig. 2), and a short vertical aerial electrically associated with one frame aerial. The output from this combination, after amplification, is arranged to operate a pen-writing oscillograph recording on a drum rotating in step with the frames. The mechanical design was to a large extent governed by available parts from the older model, and since the equipment is housed in a small wooden hut at considerable distance from trees, telegraph wires, wire fences, electric cables, etc., which might introduce lateral deviation of incoming signals, it is necessarily battery-operated.

PRINCIPLE OF OPERATION

Let us consider the recording (R) frame and its amplifiers alone. Here we have a normal frame direction-finder in which a steady stream of atmospherics of constant amplitude produces in the oscillograph coils a current which at any instant will be $I \cos \theta$, where θ is the angle between the instantaneous position of the plane of the rotating frame and the azimuth of the atmospheric stream, and I is the output current when the plane of the frame coincides with the direction of arrival; since rectified current is fed to the oscillograph the armature will always be rotated in one sense, and this rotation is arranged to produce upward deflection of the pen. The type of chart obtained will then be similar to that from the older instrument, namely one with two wide bands of record, 180° apart, with very flat maxima.

From the opposing (O) frame, rigidly fixed at 90° to the other, the same source would produce an output equal to $I \cos (\theta + 90°)$, assuming set gains to be equal and θ the same angle as formerly. This output is connected to the oscillograph coils so as to apply a torque tending to rotate the armature in a direction opposite to that resulting from the output of the R set. The use of similar amplifiers ensures that the two torques are applied to the armature simultaneously, and the resultant torque is therefore proportional to $I[\cos \theta - \cos (\theta + 90°)]$. A stop provided below the pen prevents downward deflection, so that no record will result when $\cos \theta < \cos (\theta + 90°)$.

These steps are presented graphically in Fig. 1. Curve *a* shows the usual "figure of eight" diagram, drawn to cartesian co-ordinates, resulting from the R set only; the height of the curve is a measure of the amplitude of pen deflection obtained as the frame rotates through 360°; Curve *b* shows the output from the O set (at right angles to the R set) producing pen deflection downwards; *c* combines *a* and *b*, the dotted portion below the zero line being suppressed.

Now let the signal from the non-directional aerial be added to that from the R frame before rectification and adjusted in amplitude to be equal to it when the frame is set to receive maximum signal. The output from the R set is then $I(1 + \cos \theta)$ and follows the usual cardioid diagram. This results in the disappearance of one of the recorded sectors, where $1 + \cos \theta = 0$, and the record becomes unambiguous. Curves *d* and *e* show the unrecti-

---

* Reprinted from *Journal I.E.E.*, 1938, vol. 82, p. 289.

† See Reference (1).

fied currents in the frame and vertical aerials respectively, *f* combines the two, and *g* adds the combined output to that from the O frame (*f* + *b*).

If now the amplification of the O set be progressively increased, the angle over which the output from the R set exceeds that from the O set will be correspondingly decreased. Thus the sector around the true azimuth of arrival over which a visible record is obtained may be narrowed to any desired width, and since the output from the O set is zero when that from the R set is at its maximum such reduction in sector width will not reduce the amplitude of the record at maximum, however narrow it is made. In practice a difference in gains of about 17 db. was found the most suitable value to employ, and *h* is the resultant curve with such values.

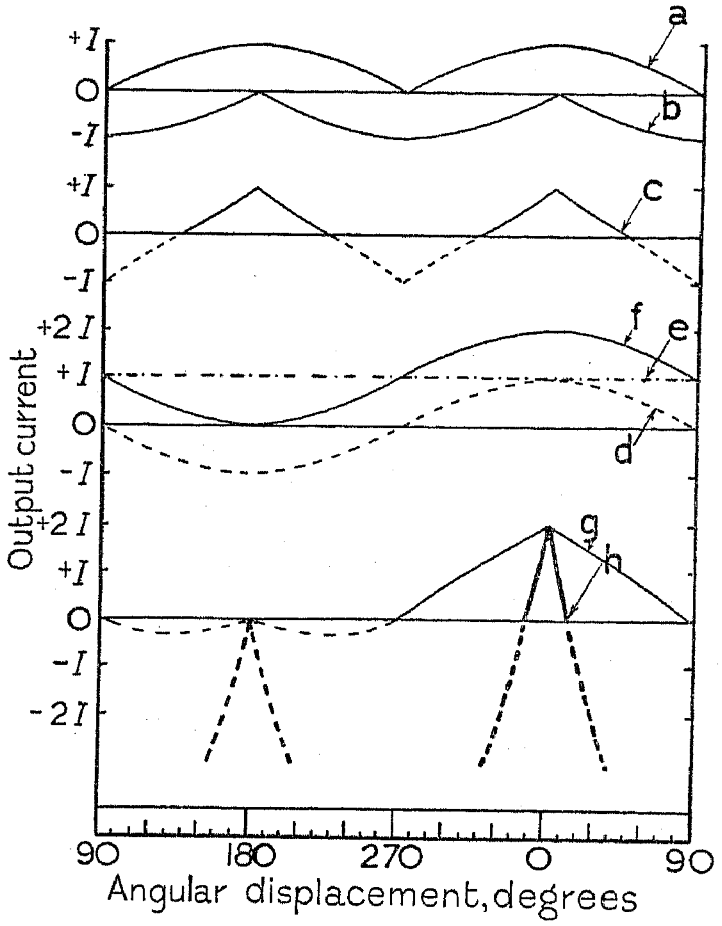

Fig. 1

Throughout these considerations it has been assumed that linear rectification is used. With the square-law rectification actually in use a further narrowing of the sector results; the output curves depart somewhat from those forming Fig. 1, but no fundamental change in the argument results.

DETAILS OF RECORDER

The general assembly of the recorder is seen in Figs. 2 and 2A. A pair of frame aerials (the recording frame R and the opposing frame O), mounted mutually at right angles on the vertical spindle A, are supported in ball-races in the timber framing B and rotated by a turret-clock movement C. The metal drum D carrying the recording paper is mechanically geared to the frame spindle, while an oscillograph E receives the amplified current from the aerials and moves a siphon pen P to record the atmospheric on the paper.

The Frame Aerials

The frame aerials, each 1 metre square, are wound in slotted paxolin rods 2·5 cm. in diameter, supported on aluminium cross-arms carried by diagonal arms of 2·5-cm. diameter aluminium tube which are secured by a central hub to the vertical spindle, of similar tube. The diagonal arms are broken near their ends, and an insulating section *f* is inserted to prevent the formation of closed conductive loops in the framing. The frames are wound with 170 turns of No. 26 S.W.G. copper wire, silk-covered, at 2·5-mm. spacing, terminated on insulators, from which leads are taken through the central spindle to the mercury slip-rings S. The centre point of each winding is similarly terminated, and the leads taken to a single point on the metal framing. The central spindle is suspended from a ball-race in the top of the teak framework B; its lower end passes through a radial race in the lower horizontal member and carries below it a flat circular porcelain disc in which are five concentric circular channels for mercury. Fixed contacts through the porcelain into each mercury " ring " terminate the four ends of the frame windings and the earth connection. Into the rings dip five iron rods mounted on an insulating block on the framework, and these are connected to screened leads which are taken to the tuning equipment.

The aerials each have an inductance of 0·034 H and a resistance of 107 ohms at 10 kc. per sec., and are adjusted to obtain zero mutual inductance.

Driving Clock

The frames are rotated four times per hour by a turret-clock movement C through a 5-mm. diameter leather belt drive. This was ultimately chosen, in view of its slight elasticity, as the best drive for producing a steady movement of the frames from the intermittent drive of the clock. A further steadying is obtained by coupling the belt pulley to the driving shaft by a spring.

Recording Drum

The hollow cylindrical drum, 15·3 cm. in diameter and 34·5 cm. long, and its supporting framework, which in the earlier instrument was connected directly to the frame spindle and mounted below the frame, is in this model supported on a bracket J (Fig. 2) attached to the side of the wooden framework and driven through a 1 : 1 bevel gearing and shaft G. Thus the drum rotates in synchronism with the frames, and any point on its surface may be referred to a particular orientation of the frame system. As the drum rotates it is lowered at the rate of 3 mm. per revolution by the screw *s* turning in the fixed nut *n*, so that a stationary pen traces out a helix on the drum surface.

The Chart

The recording chart consists of a sheet of glazed paper 49 cm. by 34 cm., secured around the drum by pasting at the 1-cm. overlap. On removal from the drum the helical track drawn by the stationary pen appears as a series of parallel and nearly horizontal lines, 3 mm. apart, and the upward deflections of the oscillograph pen caused by the atmospherics as short lines perpendicular to these.

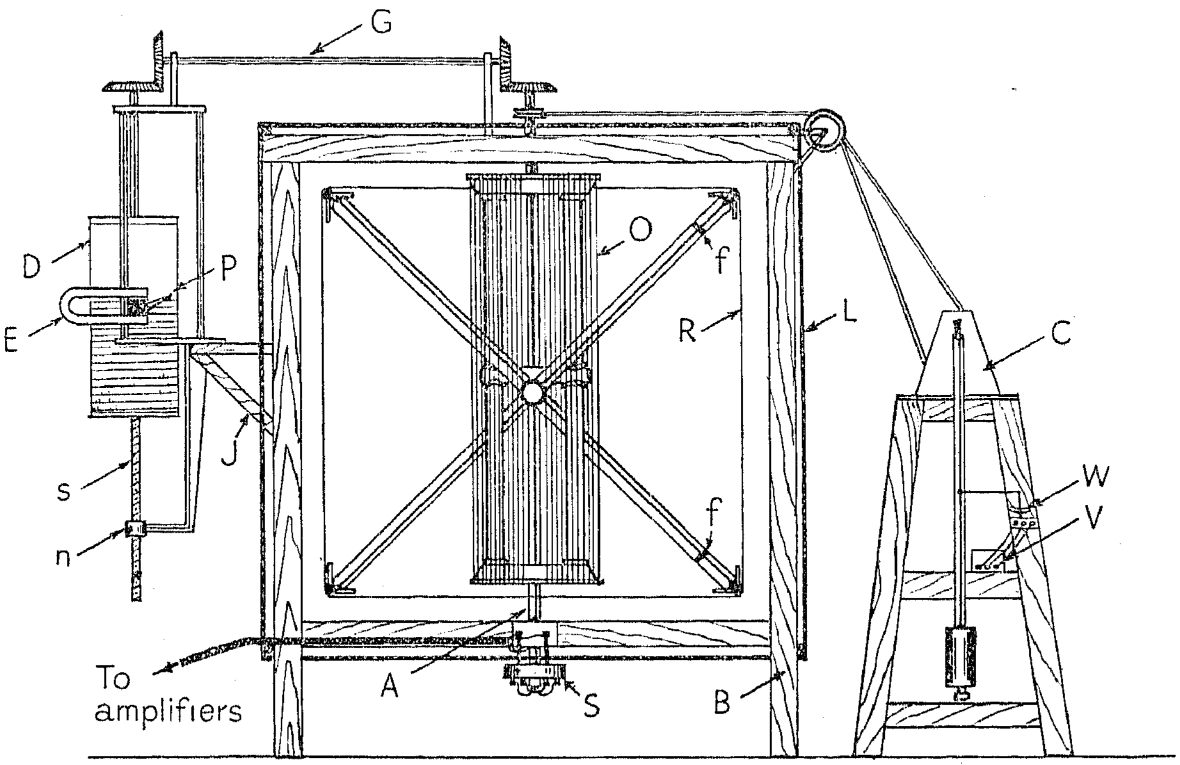

Fig. 2

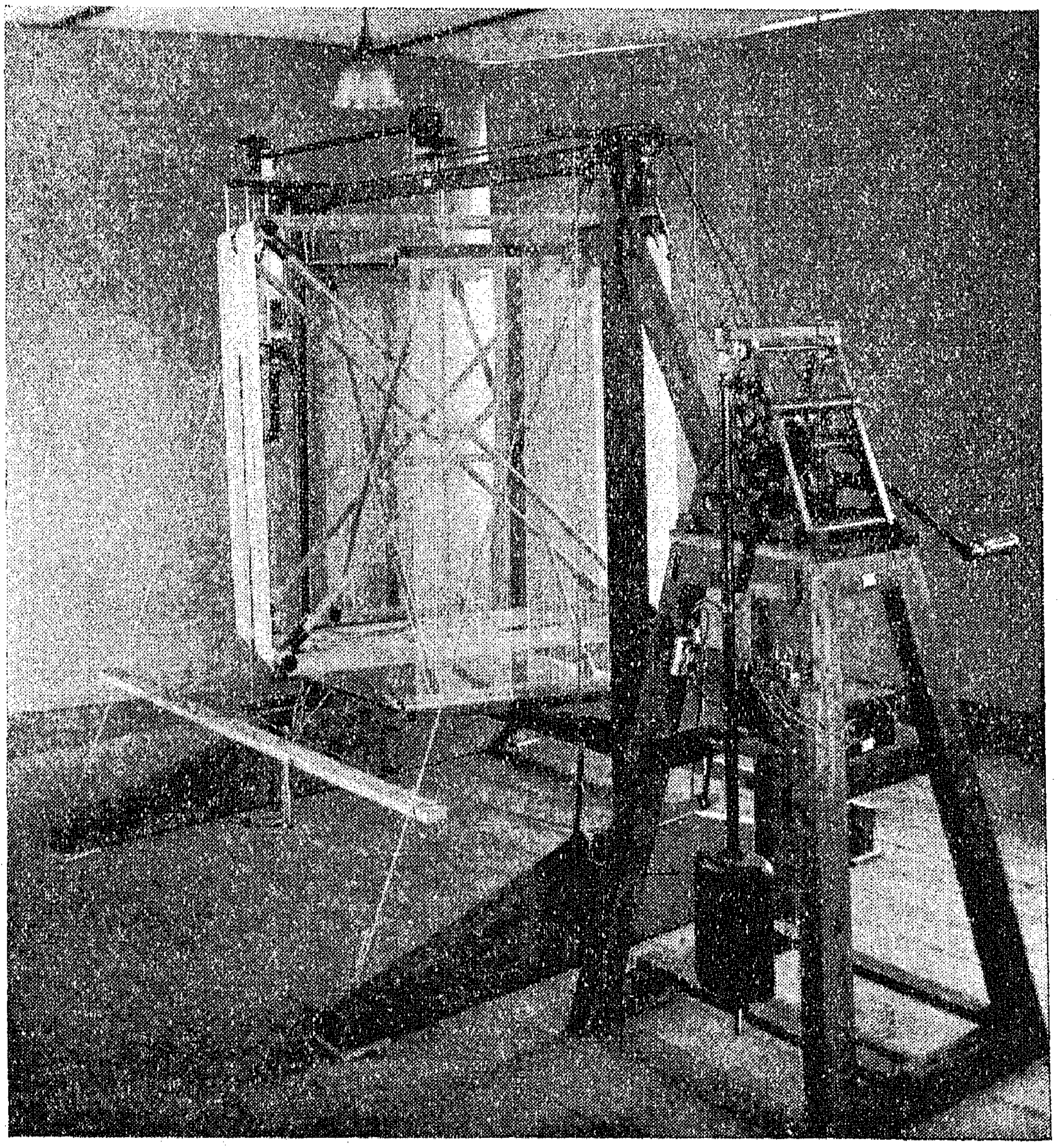

Fig. 2A

The Oscillograph

The Abraham–Bloch quadro-polar instrument with soft iron armature, used before, is again employed. It has four field windings each having a resistance of 2 500 ohms, two of these being connected in parallel in the output circuit of each amplifier, in such a manner that the flow of current in the pair connected to the amplifier associated with the R frame rotates the armature in a direction producing upward deflection of the pen, while current in the other pair from the O frame deflects the pen downwards. On the armature is mounted a light aluminium girder carrying the siphon pen. This pen consists of a short length of silver tube of 0·33 mm. bore with a writing tip of 0·12 mm. bore pushed into the larger tube. This gives a very fine line and does not tend to clog, as was found when the pen was constructed of fine bore tube throughout. A small stop fixed to the oscillograph prevents downward deflection of the pen, and the controlling torque of the oscillograph is adjusted so that in the zero position of the armature, under the steady anode current from the two sets, the pen rests lightly on this stop. To ensure uniform pen pressure on the drum surface, the oscillograph is mounted on a rotatable rod parallel to the axis of the recording drum. A small roller on the oscillograph immediately in advance of the pen is held on to the paper surface under the action of a spring-applied torque on the supporting rod, and the girder carrying the pen bent to give a certain but light contact between the pen and the chart. A small camel-hair brush is also provided to remove dust from the paper immediately in advance of the pen. This has proved very effective in preventing clogging of the pen due to dust collected from the paper surface.

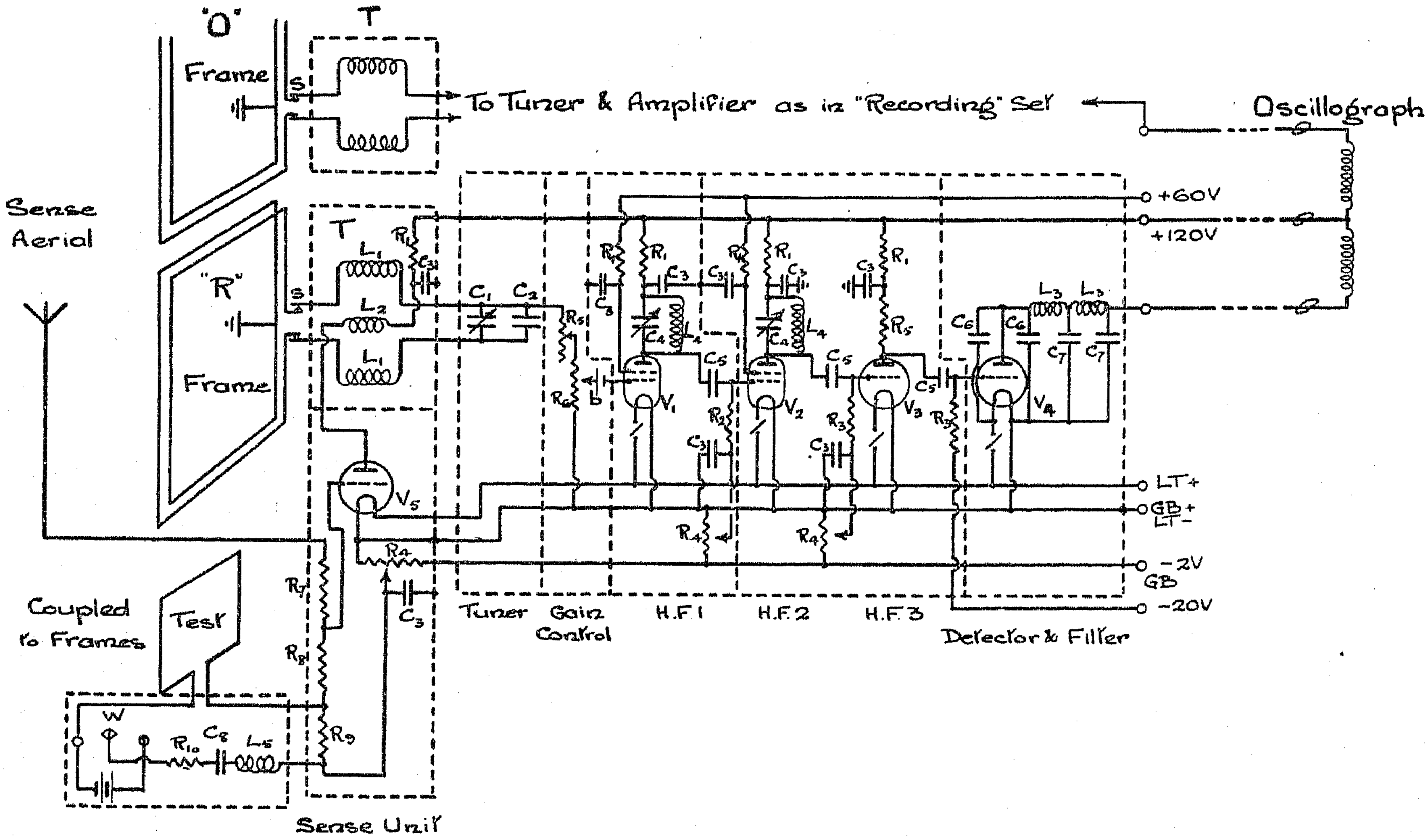

Fig. 3.—Narrow-sector recorder circuit diagram.

L.T. and H.T. batteries common to both sets. Separate grid-bias batteries for each amplifier.

R_1, 5 000 Ω.
R_2, 1 MΩ.
R_2, 0·5 MΩ.
R_4, 1 250 Ω.
R_5, 200 000 Ω.
R_6, 50 000 Ω.
R_7, 2 MΩ.
R_8, 0·1 MΩ.
R_9, 150 Ω.
R_{10}, 250 Ω.

C_1, 0·001 μF.
C_2, 10 × 0·0008 μF.
C_3, 1 μF.
C_4, 0·0013 μF.
C_5, 0·015 μF.
C_6, 0·25 μF.
C_7, 0·5 μF.
C_8, 0·5 μF.

L_1, 2 770 μH.
L_2, 88 640 μH.
L_3, 709 000 μH.
L_4, 510 000 μH.
L_5, 250 000 μH.

V_1, 215SG.
V_2, 215SG.
V_3, HL210.
V_4, P2.
V_5, H210.

Amplifiers

Similar amplifiers are provided for each of the two frames, and the schematic wiring diagram forms Fig. 3. Each has three stages of high-frequency amplification followed by a detector stage and filter unit. To ensure complete freedom from feed-back and cross-talk between amplifiers the separate units are enclosed in metal screening boxes and the connections are made throughout with screened leads. The two ends of the frame aerial are connected via the mercury slip-rings S and the split secondary L_1L_1 of the transformer T to the tuner unit. The transformer provides a means of coupling to the frame circuit the output from the vertical aerial which is discussed later. The tuner unit consists of a bank of ten fixed condensers of capacitance 0·0008 μF, inserted as required across the frames by a rotary switch and in parallel with a variable condenser of capacitance 0·001 μF. The output from the frames is taken to the grid of the first high-frequency valve, V_1, via the gain-control potentiometer. This consists of a combination of rheostat R_5 and potentiometer R_6, the former giving attenuation of the order of 0–15 db. whilst 15–40 db. is obtained on the latter. Tuned anode coupling is employed for the first two high-frequency valves and consists of two 0·33-H coils mounted astatically in a copper can and tuned by a 0·0013-μF condenser. Grid bias for the first valve is obtained from a small 2-volt cell *b*, while V_2 is biased from a potentiometer R_4 across a 2-volt cell which also feeds the potentiometer from which bias to V_3 is taken.

V_1 and V_2 are 2-volt screen-grid valves having an amplification factor of 170, while V_3, in the third high-frequency stage, is a triode. This is resistance-capacitance-coupled to the detector stage, which is a power triode with its grid biased to provide anode-bend detection.

The anode of the detector is connected to the deflecting coils of the oscillograph through a low-pass filter with a cut-off frequency of about 400 cycles per sec., and an attenuation at 10 kc. of about 80 db. This prevents the passage of radio-frequency potentials to the oscillograph and consequent re-radiation to the frames which are very close to it, a point of considerable importance in view of the high gain of the set.

Decoupling is employed throughout in both grid and anode feeds; and, as an additional precaution against cross-talk, separate bias batteries are provided for each amplifier.

The Vertical Aerial

A short vertical aerial, 6 metres high, is provided outside the building. It is connected to earth through a resistance of a few megohms so as to form an aperiodic circuit. A suitable portion of the aerial voltage is taken to the grid of valve V_5 by tapping along this resistance. V_5 is a high-impedance valve with a low-impedance winding L_2 in its anode, forming the primary of transformer T. This arrangement serves to impose the tuning of the frame circuit on that of the aerial and to ensure that the output of both shall appear correctly phased at the grid of valve V_1. By variation of the value of resistance R_7 or of the coupling between L_1 and L_2, the amplitude of the current fed to the frame circuit is adjusted to be equal to that produced in the frame itself.

Test Unit

In order that the sensitivity of the equipment may be maintained constant, it is arranged to produce artificially a transient impulse of constant amplitude, and to record on the drum the amplitude of pen deflection it produces. This test unit (V, Fig. 2) is indicated in the bottom left-hand corner of Fig. 3. A 2-way mercury switch W, operated mechanically from the pendulum of the driving clock, charges condenser C_8 from the battery and discharges it through resistance R_{10}, inductance L_5, a 4-turn loop coupled to the frames, and R_9, part of the loading resistance in the aerial. The loop consists of two turns of screened twin cable mounted in a fixed position on the main framework (see L, Fig. 2), care being taken to avoid closed loops when earthing the screen. The test signal in the loop induces in the frames signals simulating the arrival of a stream of constant-amplitude atmospherics from a constant azimuth, and the current in R_9 produces a " sense " indication in the test signal, the value of R_9 being so adjusted that the voltage appearing at the grid of V_1 from the aerial circuit exactly equals that derived inductively from loop L by the R-frame circuit with the frame set in the direction of maximum pick-up.

Since the radiation from the test loop is itself directional and fixed, the test-signal record serves as a directional reference mark as well as a gain calibration.

PERIOD OF OBSERVATIONS

Following a period of preliminary testing the recorder was taken into routine use in September, 1930, and it continued to operate as a narrow-sector recorder until May, 1933. At this time information on intensity of disturbance was desired, and the instrument was used for recording a cardioid diagram until March, 1935. A short period of narrow-sector recording followed, at a lower gain than that used previously, and recording ceased in August, 1935.

The following discussion is mainly confined to the sequence 1930–33. Throughout the period the equipment was tuned to 10 kc., this frequency being chosen to ensure freedom from signal interference without the necessity for resorting to very narrow band-width.

GENERAL APPEARANCE OF RECORDS

Four typical charts are reproduced in Figs. 4A, 4B, 4C, and 4D. The record commences at the bottom of each chart, at about noon, and progresses upward at the rate of 1 line in 15 minutes, time and angle increasing from left to right. At F, at the beginning and end of each chart, is the record made by the test unit, indicating the relative amplification and width of sector over which atmospherics are accepted by the equipment and serving also as an azimuthal reference mark. Fig. 4A is typical of records obtained during the winter period. A narrow sector of disturbance will be seen throughout the 24 hours in a direction north of west, one from the west during the night, and a third from the south-west during the early morning. Fig. 4B shows the more complex character of the record during the summer, indicating sources in widely differing directions, each source being active for a few hours. Frequently during the daytime deflections are obtained all round the chart, but azimuths in which the frequencies of occurrence reach maxima can usually be determined. Occasionally, when thunderstorms are in progress near the station, charts similar to Fig. 4C are obtained, the disturbed azimuth varying steadily with time, large angular swings often occurring within a few hours. In October and November two active channels are often visible throughout the day, one to the south, and the other to the north of west (Fig. 4D), the former attaining peak amplitude during the daytime while the latter predominates at night.

Marked differences can be seen in the character of the disturbance in the various sectors. At A in Fig. 4B long strokes of the pen occur, each stroke separated from the next by an appreciable time-interval. At D, on the other hand, considerable thickening of the baseline occurs due to repeated small pen excursions, the time-interval between them being insufficient to permit appreciable paper movement. In general, a distribution of the latter type is confined to a narrow sector, while the large individual strokes are much more scattered round the chart. In analysis it is possible to recognize four different types of distribution, and while the boundary between them is somewhat arbitrary, being assessed visually, and probably varying with the person analysing the record, it is of assistance when effects of the disturbance are being considered.

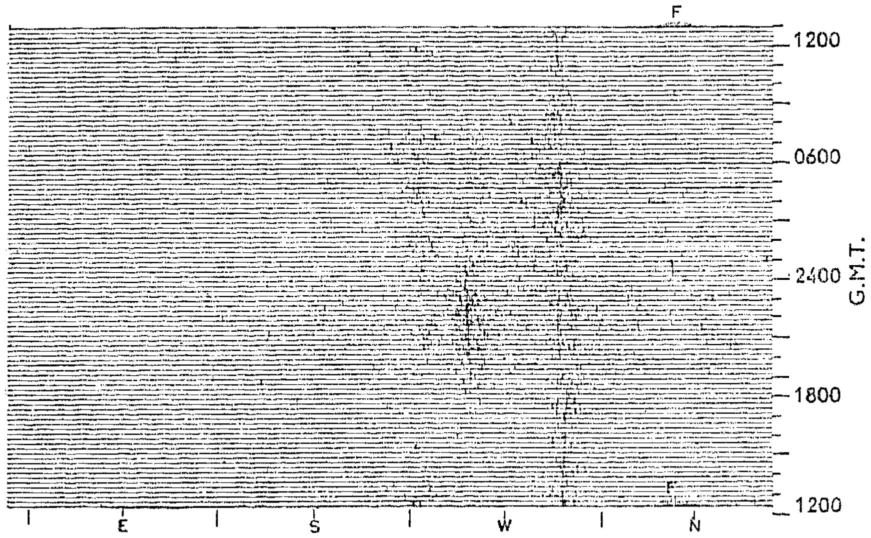

Fig. 4A.—Record for 28th–29th January, 1933.

CHARACTERISTICS OF DISTRIBUTIONS

(A) This first group consists of strokes from 4 mm. to 8 mm. in length, occurring with moderate frequency.

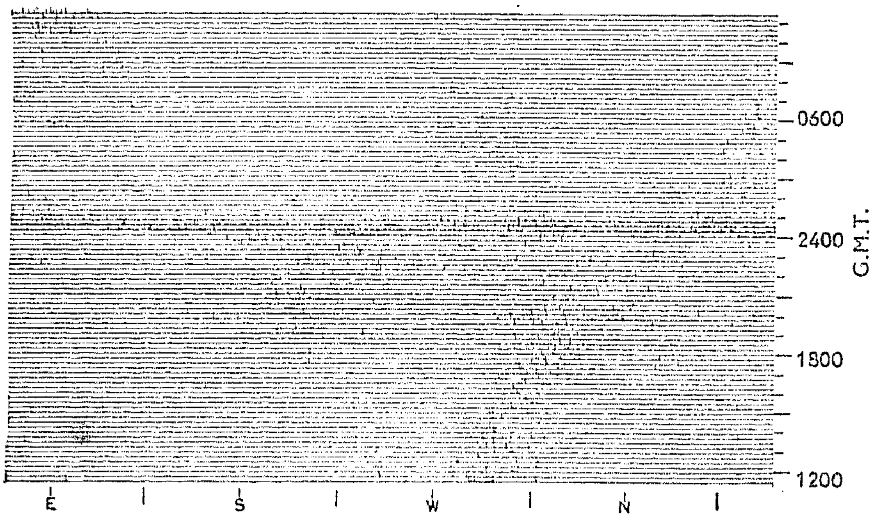

Fig. 4C.—Record for 20th–21st August, 1932.

(B) The second consists of individual strokes up to 4 mm. long occurring at a frequency at which overlapping takes place on the chart.

(C) Individual strokes up to 4 mm. long of infrequent occurrence.

(D) Overlapping individual strokes of less than 2 mm. length.

METHODS OF PRESENTATION

By placing over the original chart a transparent graticule subdivided vertically into 1-hour time divisions and horizontally into 10° steps, the disturbed azimuth and types of activity present at any hour were abstracted. This information was then carried to monthly charts, a letter for each day of the month being used and inserted in the appropriate " azimuth-hour " block indicating the hour at which a particular azimuth was disturbed. Characteristic colours for the letters were further employed to indicate the type of disturbance as outlined above. Figs. 5A and 5B show the charts for December, 1932, and August, 1931; for reproduction in monochrome the daily letter has been dropped and characteristic symbols have been adopted for the four types of distribution, namely a long line for Type A, a large dot for B, a short line for C, and a small dot for D.

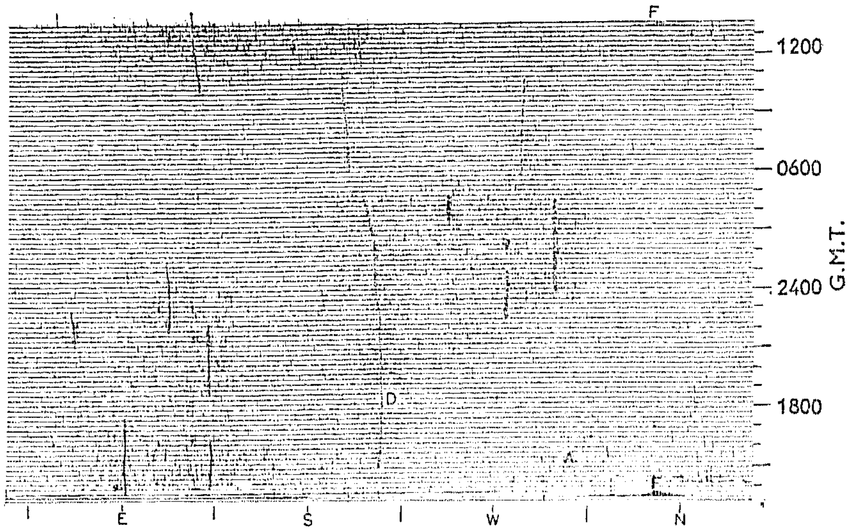

Fig. 4B.—Record for 5th–6th May, 1931.

These monthly charts have been collected into annual charts; that for the period August, 1931, to July, 1932, forms Fig. 6. Here the number of dots in a block 5° (horizontally) by 1 hour (vertically) indicate the number of days in the year on which that particular azimuth was disturbed at that hour.

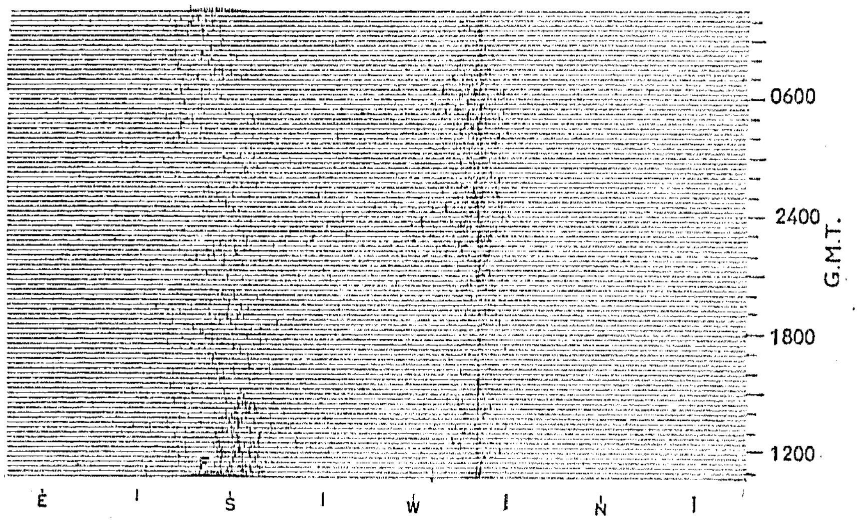

Fig. 4D.—Record for 19th–20th October, 1930.

In view of the difficulty of publishing the 12 monthly charts complete, three periods have been taken, each covering 3 hours, centred on the hours 1000, 1600, and 2400 G.M.T., these periods showing the general trend of the diurnal variation of disturbed azimuth. Curves have been prepared in which the frequency with which any azimuth is disturbed in any month is shown by the height of the curve, the frequency of disturbance being expressed as a percentage of the total number of days in the month (Fig. 7). In addition to indicating the frequency, the area below the curve has been shaded to indicate roughly the mean intensity of the disturbance from each sector. This is discussed later.

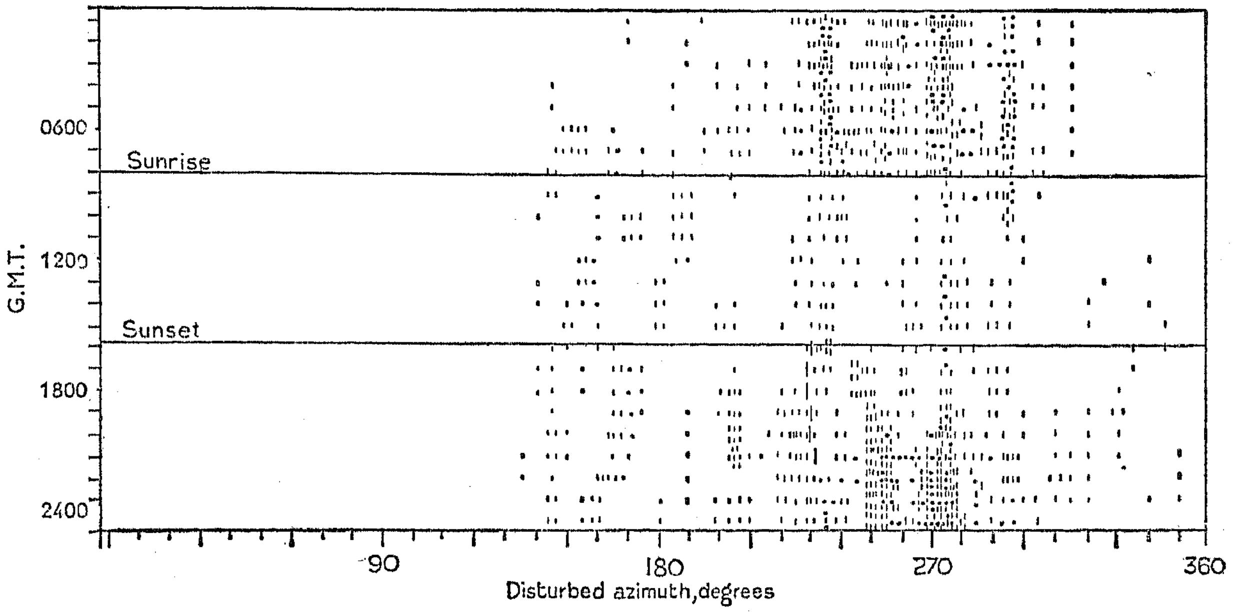

Fig. 5A.—Distribution of disturbed azimuths at each hour, December, 1932.

| Type A
• Type B
ı Type C
· Type D
} Disturbance characteristics.

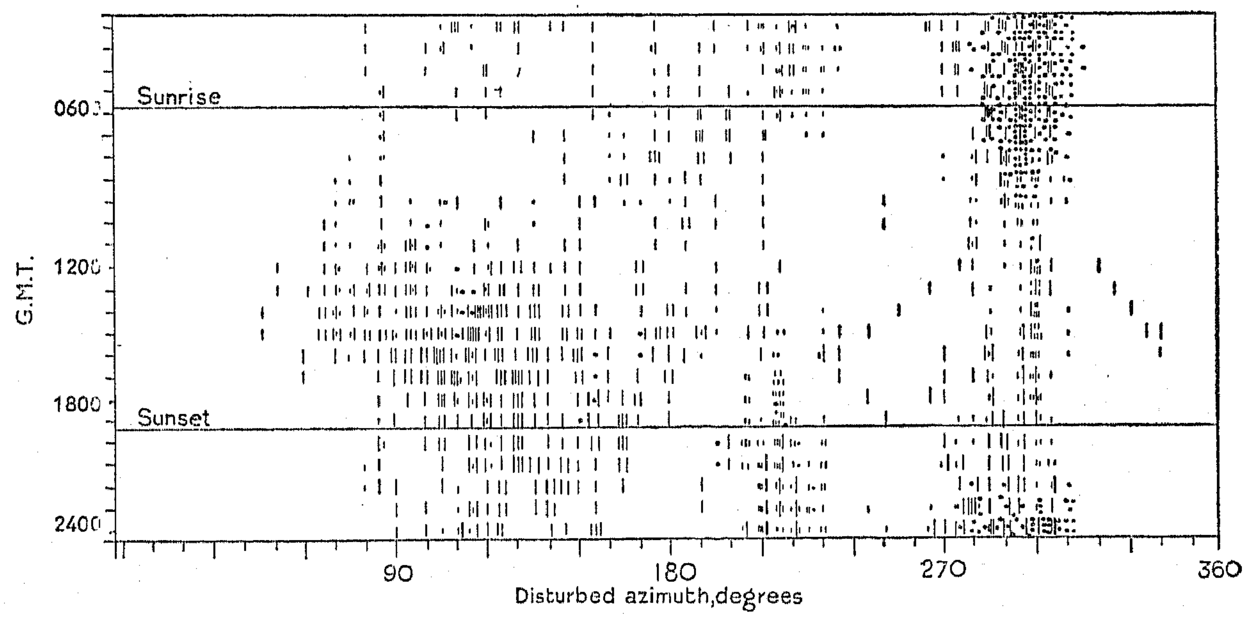

Fig. 5B.—Distribution of disturbed azimuths at each hour, August, 1931.

| Type A
• Type B
ı Type C
· Type D
} Disturbance characteristics.

DISTURBED AZIMUTHS: DIURNAL AND SEASONAL VARIATIONS

Considerable variation in the disturbed azimuth occurs both diurnally and seasonally, some variations being systematic whilst others appear to be of a purely random character.

Summer

Fig. 4B shows a normal day chart, and Fig. 5B a mean monthly distribution diagram for the summer months. During the afternoon, disturbance arrives from the east and south-east and is of Type A, large individual disturbances occurring at moderate frequency. This afternoon disturbance is generally scattered over the sector 90°–200°, but the directions of peak intensity are alone plotted on the diagram. The manner in which these directions of maximum intensity vary from day to day over the quadrant will be seen in Fig. 5B. Activity commences at about 1000 G.M.T., is at its peak in the afternoon, and continues well into the evening hours.

An hour or two after sunset, sources to the west and north-west become the prominent feature. Individual disturbances are less intense than those obtained during the day, but the frequency of occurrence is much higher. The azimuth of maximum does not vary from day to day in the manner of the afternoon disturbance, nor is so wide a sector disturbed. Maxima lie between 270° and 300° during the summer. Whilst activity commences some 2 or 3 hours after sunset, it dies away at about 0700 or 0800 G.M.T., irrespective of the time of sunrise.

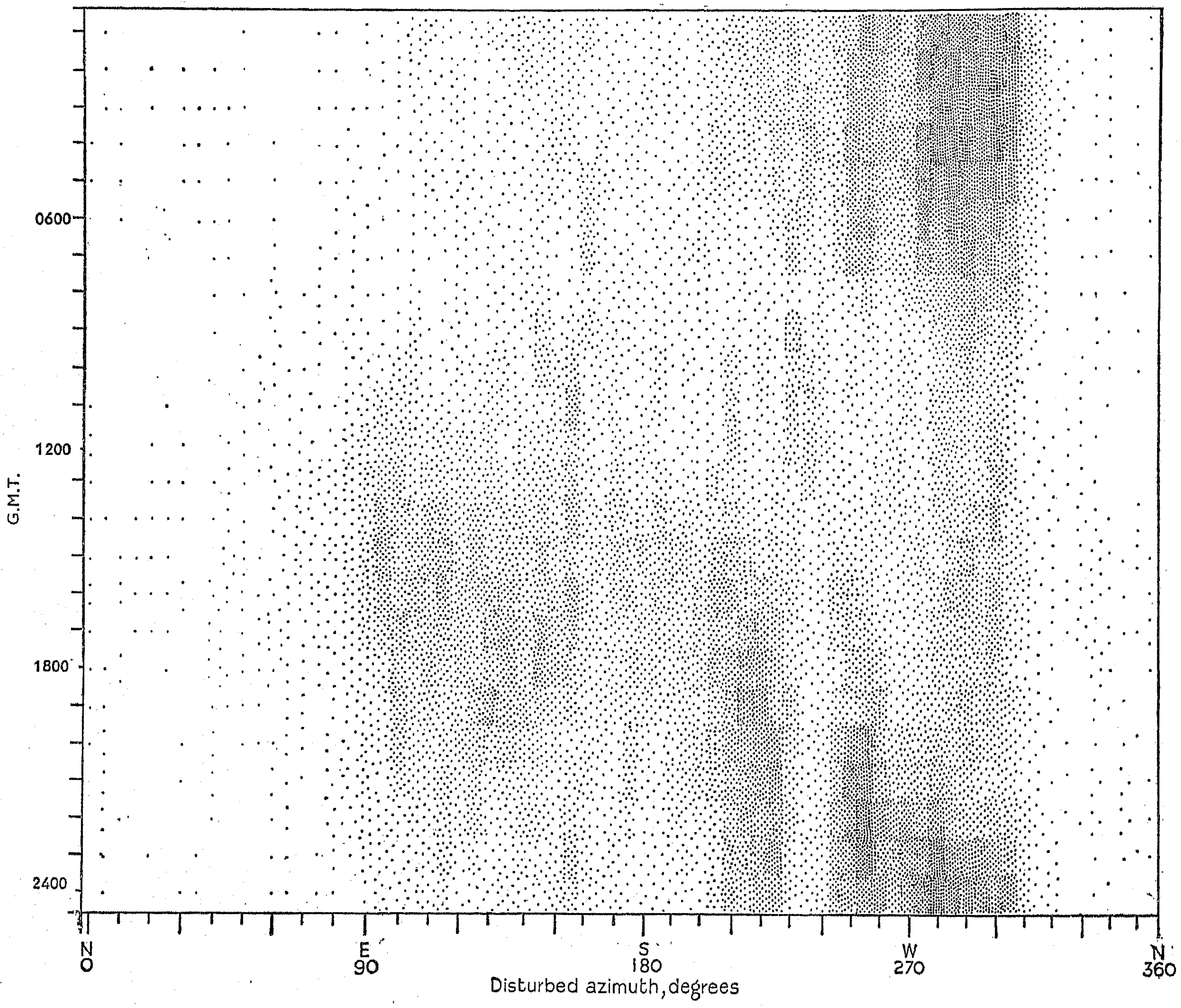

Fig. 6.—Annual distribution of disturbed azimuths at each hour.

During the evening and throughout the night hours another active channel, at 210°–230°, is often shown. The intensity of individual disturbances is relatively small, but often the frequency of occurrence is quite high. Maximum activity in this direction occurs during April and May.

Winter

During the winter months very little disturbance arrives from the east or south-east. Occasional large individual disturbances are registered in this direction, but the most prominent source at all hours of the day lies to the west (Figs. 4A and 5A). Four distinct sectors are found, centred on 235°, 250°, 275°, and 290°. During the day activity is usually of Type C, but an hour or two after sunset the received intensity and also the frequency of occurrence increase rapidly, and a Type B distribution obtains. Maximum activity occurs at 250° and 275° during the hour around midnight, while at 235° and 290° it occurs after midnight and continues until early morning. The four centres are not often found on a single chart, and the relative amplitude of disturbance in each sector varies from day to day. On the record reproduced as Fig. 4A the sector at 290° contains the heaviest disturbance, while practically nothing is found at 270°. On this record a little activity (Type C) will be seen to the south during the evening, but this seldom attains an amplitude comparable with that from the west nor is the frequency of occurrence so high except in the months of October and November (Fig. 4D), when southerly activity is sometimes found throughout the day, with its peak in the afternoon. As summer approaches, the sources at 270°

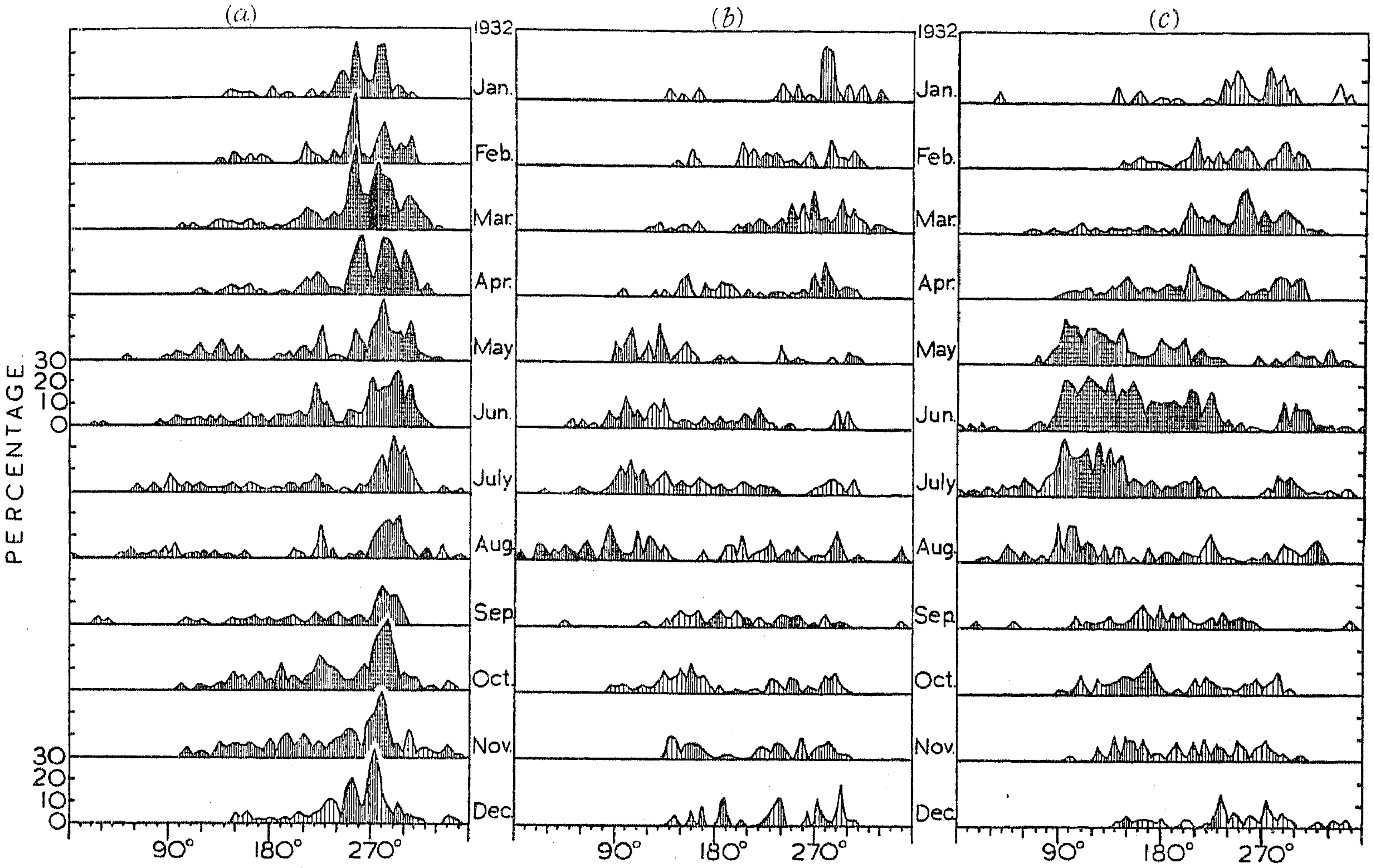

Fig. 7.—Percentage of month with disturbance at selected hours in sectors 5° wide.
(*a*) 2400 G.M.T. (*b*) 1000 G.M.T. (*c*) 1600 G.M.T.

Key to intensity shading: ▥ 1 ▥ 2 ▦ 3 ▩ 4 ■ 5

and 290° spread northward, while the sector found in December at 235° disappears and activity around 210° becomes prominent during the early evening.

The variations from month to month on these general tendencies may be seen in Fig. 7. Points which will be noted are: (1) the persistence of the westerly sources throughout the year during night hours, (2) the northerly movement of the night-time sources with the approach of summer, (3) the similarity between the azimuthal distributions in the forenoon and in the afternoon, (4) the increase of activity in these channels during the afternoon, (5) the limitation of the disturbance from south-easterly sources to the summer season.

The annual distribution diagram (Fig. 6) shows the number of occasions on which a particular azimuth was disturbed at each hour throughout the year. This number may be obtained by counting the number of individual spots in a block, 5° wide by 1 hour high. The scattered character of the afternoon disturbance and the concentration in azimuth of the night sources are clearly shown. Other points of interest which will be discussed later are: (1) the almost complete absence of activity from the northern quadrant extending from about 310° to 70°, (2) the narrow channel at 235° with very little activity, (3) the reduced activity around 265°, (4) the few occasions on which disturbance has arrived from the south.

This chart, although for one year only, may be taken as typical of those obtained for the two complete years and portions of others covered by this analysis. How closely the distribution is repeated from year to year may be gathered from the annual curves for 1931 and 1932 respectively, reproduced as Fig. 8. The general trends of the two are very similar, but minor variations will be seen in the positions of the peaks. These variations are relatively unimportant to the radio engineer, but to the geophysicist they may be of considerable interest when

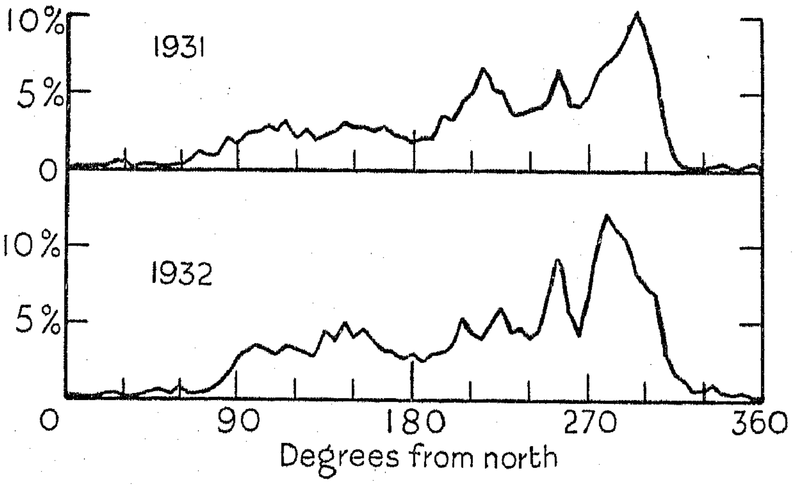

Fig. 8.—Mean annual frequency of disturbance, showing the percentage of the whole years during which disturbance was registered in any sector 5° wide.

variations from year to year in the regions of maximum thunderstorm activity and variations in propagation conditions are under consideration.

INTENSITY OF DISTURBANCE

The variation in intensity of the disturbance cannot be accurately obtained from the instrument described in this paper, for the following reasons.

Suppose the frames to be in the position to record the disturbance in the narrow sector about angle θ. If at the same instant as an atmospheric arrives along this channel a second atmospheric is received of bearing $\theta \pm (>20^\circ)$, that atmospheric will produce a current in the opposing set and will prevent registration of the atmospheric at θ°. In the limit, given two channels of disturbance 90° apart and atmospherics arriving along each with a time-interval between them shorter than the train of oscillations generated in the receivers by each, then no record would be obtained on the chart at all. Such extreme conditions have never yet been experienced in Great Britain, but it is probable that some few individual disturbances will be lost on this account and therefore too close an interpretation of intensity from records from this type of instrument must be avoided.

The provision of a switch to cut out the opposing set for alternate revolutions of the frame was at one time considered, to provide alternate lines of "cardioid" and "narrow sector" record; but, in view of the very adequate intensity analysis already covered by the earlier paper and the confusion that might arise in analysis, such a switch was never used. Some indications of the intensity of the individual disturbances which are registered may not, however, be out of place here.

The amplitude of pen deflection due to the atmospherics received over a short period was compared with the deflection obtained at the same time on the cathode-ray direction-finder,[*] tuned to the same wavelength, these deflections being recorded on film for detailed comparison. This apparatus had been calibrated in terms of the amplitude of a deflection produced by a steady field of measurable intensity. The large individual disturbances of Type A, appearing as a line of 6 or 7 mm., were found to correspond to a field in excess of 10 millivolts per metre, those of the second and third groups of 5 millivolts per metre, while those of the fourth represented a field of less than 1 millivolt per metre, the smallest recordable deflection being of the order of 0·25 millivolt per metre.

While the amplitudes of the individual disturbances of Type A may be quite large, to the radio engineer they may not be so troublesome as those of Types B or D because of the greater frequency of occurrence of the latter. To give some indication of the "disturbance value" at any azimuth, shading has been added to Fig. 7. This value was determined visually from the original charts by matching the number and intensity of individual disturbances obtained in a 10°-by-1-hour block on the chart against a code card, graded from 0 to 10. No. 1 of the code would be given when 5 disturbances per minute, each not exceeding 1 millivolt per metre, were registered; No. 2 would be used for 15 per minute each not greater than 1 millivolt per metre, or 3 per minute each not greater than 5 millivolts per metre; and so on, until No. 10 labelled a record completely inked over—in which neither frequency of occurrence nor individual intensity could be measured. Occasionally No. 7 of the code had to be used, indicating a frequency of some 200 per minute, many attaining at least 5 millivolts per minute. The mean hourly values for each month were used to determine the depth of shading in accordance with the key below the Figure.

SOURCES OF DISTURBANCE

From the earliest days of radio communication it has been recognized that visible lightning produces the prolonged "crashes" heard in receivers, and many workers, among the earliest being Popoff[*] and Fenyi,[†] have investigated the possibility of forecasting by radio the approach of thunderstorms. There are, however, other types of atmospherics, referred to in the literature as "grinders," "rattle," etc., appearing on our records as distributions of Types B, C, or D, the origin of which has been a matter of controversy. De Groot[‡] suggested that they were produced by the bombardment of the higher regions of the atmosphere by cosmic rays, and estimated the height of the region as from 200 to 600 km. In his early work Schindelhauer[§] supported this view of their origin, and when it was shown by Watson Watt,[||] in the work from which the present investigations developed, that they had marked directional properties Schindelhauer[¶] introduced the effect of the earth's magnetic field, and its action in producing a spiral motion of the incoming particles, to explain these directional characteristics and their variation with time and season. Hubert,[**] from observations in West Africa, concluded that they were due to invisible discharges occurring at an altitude of some kilometres, at tens of kilometres from the receiver, between air masses of differing potentials, while Austin[††] suggested that discharges between the meteorological cloud system over Mexico and Central America, and the electrified layers above it, produced the static of a "grinding" character heard in the U.S.A. All these explanations imply that the source of the static is local to the receiver, certainly within a few hundred kilometres. On the other hand, Bäumler[‡‡] suggested that since he obtained atmospherics simultaneously from stations very far apart, the source must be at a great distance relative to that between the separate receivers. Measurements by Wilson[§§] of the energy dissipated in a lightning flash showed that radiation from it could be expected to affect receivers over most of the earth, and led both Appleton and Watson Watt to conclude that tropical thunderstorm regions were probably the origin of the frequent low-amplitude atmospherics under discussion. These two workers differed, however, in their opinions as to the extent to which the visible flash could be regarded as the sole source of the atmospheric. Appleton[||||] maintained that the origin of all atmospherics was to be found in lightning flashes, and in a recent publication[¶¶] he reiterates this point of view, while the work of Watson Watt[***] on locating the source by audio direction-finding at a number of stations had shown that many atmospherics originate in rain regions where no lightning flashes had been recorded by the meteorological observers. Further support was given to the theory of tropical thunderstorm origin by the short series of observations made by Round, Eckersley, Tremellen, and Lunnon* in their voyage to Australia and back. The estimate that some 100 flashes occur each second over the whole globe, made by Brooks,† and the assumption that a large number of the atmospherics generated by them would affect receivers in all parts of the world, indicate that atmospherics in sufficiently rapid succession to account for the " rattle " or " grinding " noise may reasonably be expected from such sources, and support the hypothesis of tropical thunderstorm origin, leading to an examination of the bearings of the atmospherics recorded in relation to the bearings of regions of tropical thunderstorms.

---

[*] See References (2) and (3).

[*] See Reference (9).
[†] *Ibid.*, (10).
[‡] *Ibid.*, (11).
[§] *Ibid.*, (12).
[||] *Ibid.*, (13).
[¶] *Ibid.*, (14).
[**] *Ibid.*, (15).
[††] *Ibid.*, (16).
[‡‡] *Ibid.*, (17).
[§§] *Ibid.*, (18).
[||||] *Ibid.*, (19).
[¶¶] *Ibid.*, (20).
[***] *Ibid.*, (21).

In Figs. 9A and 9B the results of the thunderstorm investigation made some years ago by Brooks have been set out on a zenithal azimuthal graticule centred on London, produced by J. St. Vincent Pletts. This map projection shows the true distance and direction from its centre of any point on the earth's surface; thus the azimuths of the several thunderstorm centres and their relative distances from the receiver are clearly presented, and the geographical nature of the paths to the receiver is easily realized. The shaded areas on the two charts show regions where storms occur on more than 20 % of the days during the periods April–September (summer) and October–March (winter). It will be seen that three active regions exist, one in the East Indies, another in Central Africa, and the third in Equatorial America, bearing 60°–80°, 140°–170°, and 210°–300° respectively. In all of these regions thunderstorms were found to occur on at least 30 % of the days of the year, and if we consider the hours of peak activity in any region to be during the late afternoon (local time) we should expect maximum disturbance to be registered here from each of the centres at about 0600–0800, 1400–1600, and 1900–2100 hours G.M.T. respectively.

To enable direct comparison to be made, curves showing the frequency with which disturbance was registered in any direction are inscribed within the boundaries of the charts. In Fig. 9A, curves for the winter periods 1930–31 (Curve *a*), 1931–32 (Curve *b*), and 1932–33 (Curve *c*), are drawn and averaged to form Curve *d*, whilst in Fig. 9B those for the summer of 1931 (Curve *a*) and the summer of 1932 (Curve *b*), with *c* as mean, are shown. Thus we may read off directly that the most frequently disturbed sector in 1931 lay on the great circle passing through the east-coast regions of North America and Mexico, and that the 5° sector centred on this azimuth was disturbed for 12 % of that summer period; the amplitude of the curve is measured radially inwards from the boundary in accordance with the associated scale indicating the percentage of the period during which atmospherics were registered in any sector 5° wide.

There can be little question that disturbance in the channel 250°–280°, the prominent feature of records during the night hours (see also Fig. 7), originates in the American thunderstorm belt. The mid-winter peaks at 255° and 275° respectively indicate sources in the north-east corner of South America and the Mexican Gulf. With the approach of the northern summer the 280° channel spreads northward to include the eastern coastal regions of North America, while the 250° South American source wanes, thus fitting closely the northward movement of the thunderstorm belt with movement of the sun to its northern solstice. These westerly sources gradually die out between about 0800 and 0900 G.M.T. irrespective of season, this being much later than one would anticipate for the cessation of storms on the American continent, but their appearance in the evening appears to lag some 2 hours behind the time of sunset, suggesting considerable control of the received amplitude by transmission conditions across the Atlantic.

The summer afternoon activity, from the east and south-east (Fig. 7), characterized by its individual disturbances of large amplitude and producing the peaks in this direction shown by the curves of Fig. 9B, undoubtedly originates in the European thunderstorm regions. The increase in April, the peak in June, and the decrease in autumn, all follow closely the increase and decrease in solar altitude over this area. The angular variations of the azimuth of peak disturbance suggest moving sources at moderate distances, and the earlier work with several recorders of the old pattern has shown how the movements of the sources may be followed by two or three recorders suitably located.* Infrequently-occurring large individual disturbances, recorded over this sector during the winter months, have their origin in the discharges which even at this season accompany the passage of cold fronts over Western Europe.

The sector 200°–220°, activity in which appears during the late afternoon and evening in most months, would include the west coastal regions of France, Spain, North Africa, and the north-east coast of South America.

Very seldom are atmospherics recorded from the sector 330°–60°. In summer, a thunderstorm centre passing over the northern part of the British Isles produces a little disturbance, but in general this freedom agrees with the paucity of thunderstorm activity in this region, as is to be expected. An occasion when disturbance occurred in this sector has already been mentioned (Fig. 4C). Here the main source was a large storm centre passing across the Midlands from west to east, while a secondary centre moved up from the south-west, passed over the receiver producing the omni-azimuthal record at about 2300–2400 G.M.T., passed away in a north-westerly direction, and died out. Contrary to expectation are (*a*) the almost complete absence of disturbance at 60°–70°, the sector including the East Indian thunderstorm region which is active during the early morning hours here; and (*b*) the few occasions on which the bearing 150°–170° reveals the rapid succession of discharges known to be occurring in Central Africa, particularly during our winter. In analysis of the records from the ambiguous recorder† it was the practice to interpret continuous disturbance recorded at 60°–100° or 240°–280° throughout the night and early-morning hours, as indicating sources to the south-west at night merging into those from the north-east in the early morning as the solar influence in the Far East became effective in thunderstorm production. It is now quite certain that nothing reaches this country from the east with intensity greater than 0·2 millivolt per metre, while from the west intensities 10 times as great are normal. It is improbable that this difference is due to varying amplitude at the source, so that the explanation must be sought in the difference in transmission conditions over the paths travelled in the two cases.

---

* See Reference (21). † *Ibid.*, (4).

* See References (1), (5), and (6).
† See Reference (1).

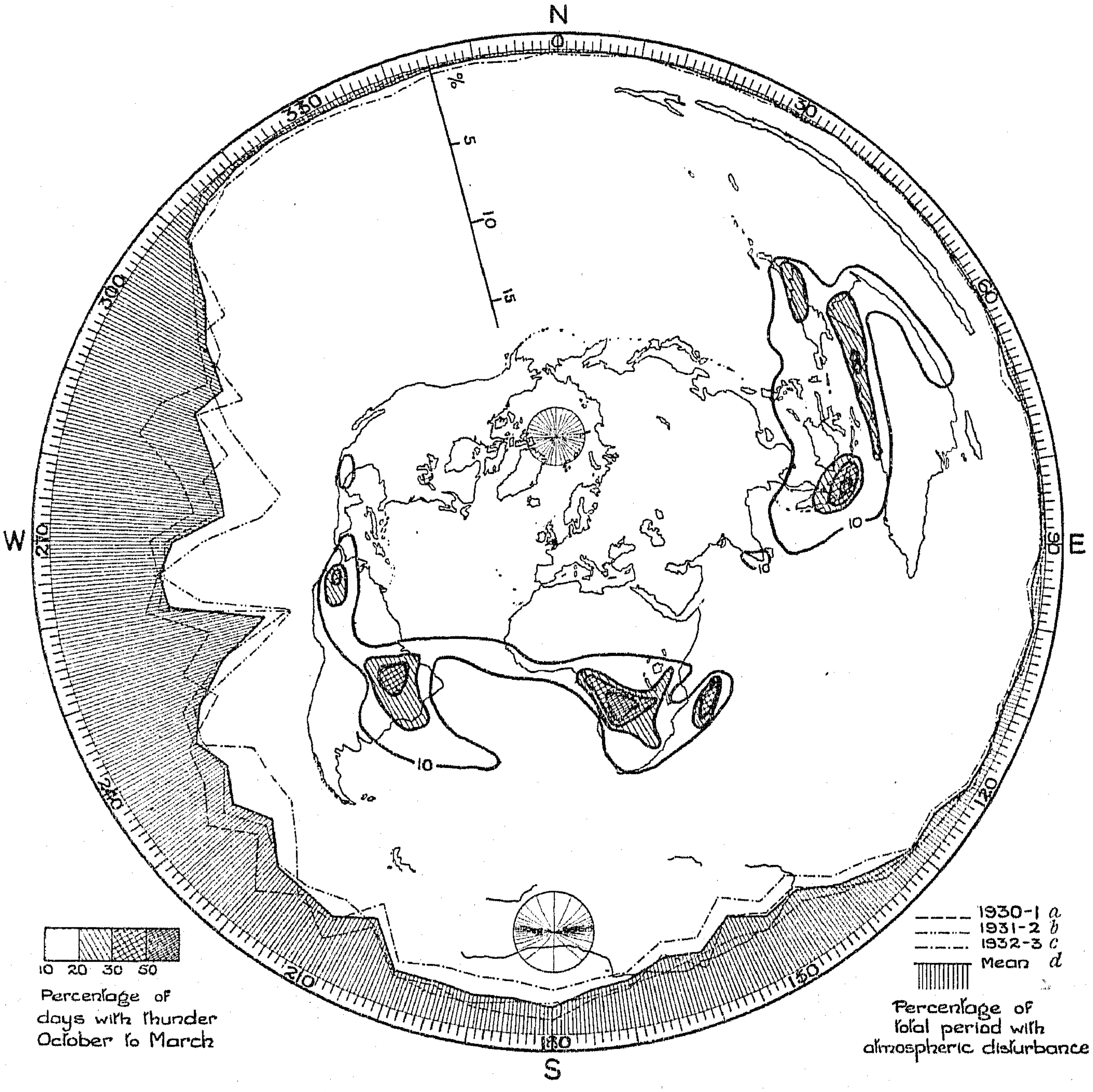
N
S
E
W
0
30
60
90
120
150
180
210
240
270
300
330
%
5
10
15
10
10
10
10 20 30 50
Percentage of days with thunder October to March
1930-1 a
1931-2 b
1932-3 c
Mean d
Percentage of total period with atmospheric disturbance

Fig. 9A

The path to the East Indies is almost completely overland, crossing Europe, Central Asia, and Indo-China; whilst that from the West Indies and South America is almost entirely a sea path. At times when the sources are active in the east the land path is in full daylight, whilst those to the west are productive when a dark path obtains. The distances in the two cases are similar, so that if we assume comparable energy at the source the difference in received intensity, a change of at least 10 : 1, must be produced by the variation in (*a*) ground conductivity and (*b*) night-time and daytime attenuation in the medium. That (*b*) is of real significance even at these frequencies is suggested by the manner in which the westerly source is suppressed until some 2 or 3 hours after sunset.

The Central African source has also to traverse a daylight land path, with the result that while occasionally southerly disturbance has the characteristics one would expect from such a source, it never attains the intensity reached by that from the west.

This absence of atmospheric disturbance from the East Indian and African sources has been also found by Bureau, working with a similar type of recorder in Paris in 1933 and 1934;* while Munro, Wark, and Higgs,† working in Australia, find these areas the sources of their maximum activity and obtain little disturbance from the American centres. From the geographical position of the Australian receiver there is a sea path and comparable distance to both African and American centres. Thus most of the variation in received intensity in their case is attributable to ionization conditions above the earth and not to surface conductivity.

It has been stated earlier that during the northern summer the western sources have a much more northerly origin. From this it must not be inferred that during the winter months no disturbance arrives from North

---

* See Reference (7).

† *Ibid.*, (8).

America. Reference to Fig. 9A shows the peak around 280° in 1931–32 to be as high as that at 250°, although from Brooks' analysis thunderstorms occur on less than 10 % of the days in the North American area. In 1930–31 and 1932–33 this peak at 280° was less prominent, and this would appear to indicate variation in the relative intensity of thunderstorm activity in these different parts of America from year to year. How accurately these curves can be used as an indication of these variations cannot be stated without day-to-day comparison with world-wide meteorology. Were the meteorological data adequate such a comparison would be economically prohibitive; but, to test how accurately the thunderstorm variations in North America are being indicated at present, a comparison was carried out for January, 1933.

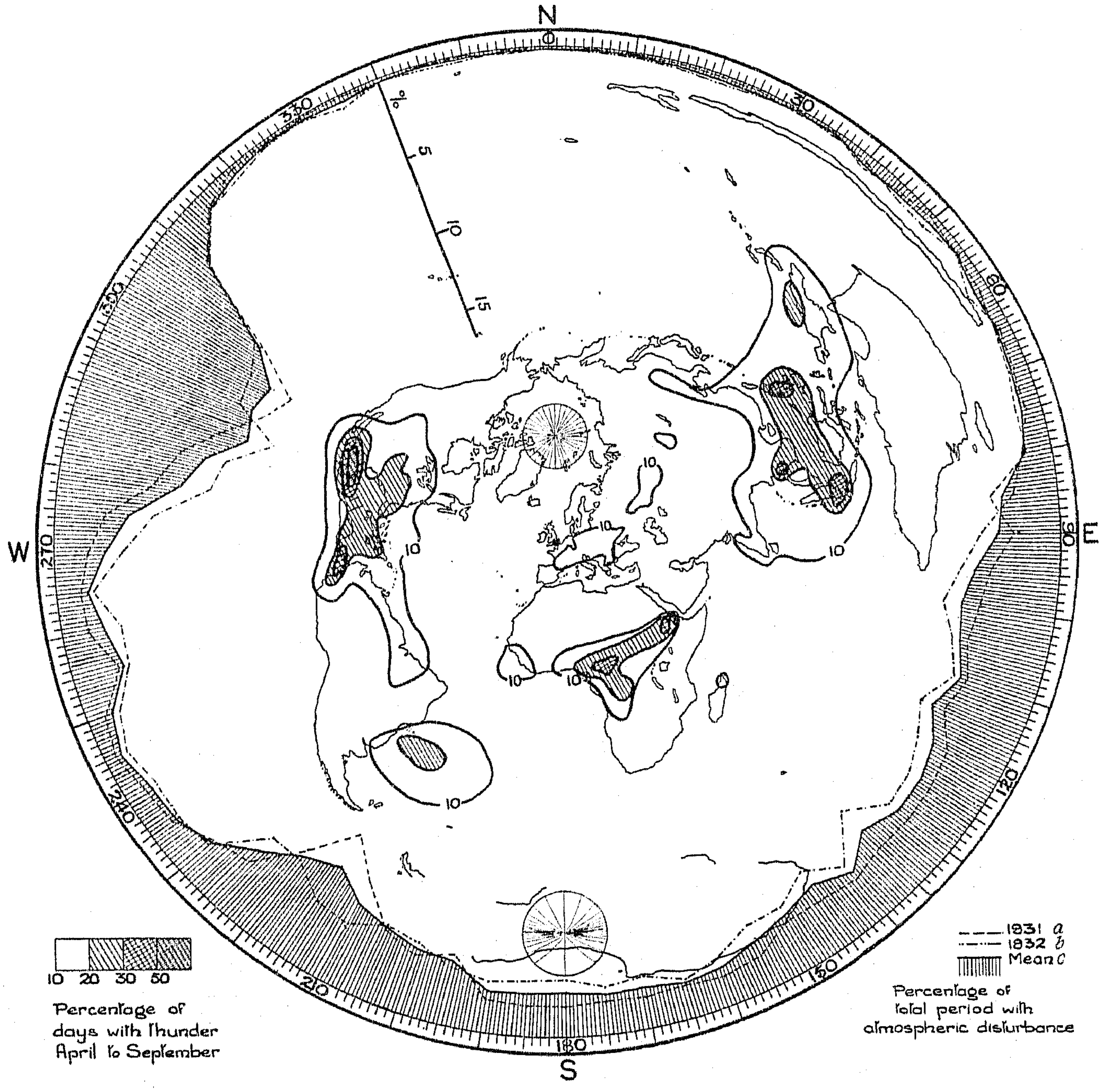
N
S
E
W
0
30
60
90
120
150
180
210
240
270
300
330
0%
5
10
15
10
10 20 30 50
Percentage of days with thunder April to September
1931 a
1932 b
Mean C
Percentage of total period with atmospheric disturbance

Fig. 9B

The recorder charts indicated a few hours' activity in this channel on each of the first 5 days of the month. It was then almost undisturbed until the 17th, except for 1 hour during the night of the 7th–8th and for 7 hours during the night of the 11th–12th. Considerable disturbance was then registered from the 17th until the 22nd, a quiet period from the 22nd to the 24th, and then further disturbance at night until the end of the month, with activity during the daytime also during the last 4 days.

During the early part of the month a succession of shallow depressions crossed the Northern States and Canada, with practically no evidence of associated thunderstorm activity on the Continent. The activity registered during the first few days probably originated in the squall conditions with heavy hailstorms in the Northern Atlantic at that time, a thunderstorm being reported off Nova Scotia on the 1st. The two nights with activity during the otherwise quiet period 6th–17th coincide with recorded thunderstorms near Boston on the 7th and in the Southern States on the 11th.

On the 16th a depression moved in from the Pacific, followed by a second on the 18th, and crossed the country, resulting in widespread thunderstorms in the Central States and Mississippi Valley on the 18th, 19th, and 20th–22nd, coinciding exactly with periods when atmospherics were recorded.

Two further depressions originated in the United States about the 24th, rapidly deepened, and had moved out to sea off the east coast by the 27th. There they remained stationary, still deepening, for some days, and finally filled up and passed northward at the end of the month. These were associated with widespread thunderstorms in the South-eastern States on the 24th–25th, but thereafter only one or two stations reported thunderstorms each day. The atmospheric records suggest much more intense electrical activity during the days when the associated fronts were over the Atlantic than during their travel over land; whether the discharge originating the atmospherics was visible lightning which escaped direct observation at sea, or invisible inter-cloud discharge, must remain a matter for conjecture. Very high winds were associated with this centre, reaching hurricane force over sea on the 26th and 27th. There seems little doubt, however, that the sources of the atmospherics were associated with this intense depression, and it is not without interest that during the time that it was active over sea atmospherics were received during the whole 24 hours, although during the daytime the intensity of disturbance was much less than at night.

It would appear then that the majority of the atmospherics producing disturbance in long-wave reception in this country are generated in distant regions of thunderstorm activity. They suffer variable attenuation in propagation, depending on the ionization conditions of the path along which they travel, with the result that American sources supply the majority of those with a field strength exceeding 1 millivolt per metre in England. Summer afternoon and evening sources lie in the thunderstorms of Europe, but American sources always predominate during the night and early morning.

ACKNOWLEDGMENTS

The work described above was carried out under the direction of Mr. R. A. Watson Watt in connection with the programme of the Radio Research Board, and the description is published by permission of the Department of Scientific and Industrial Research. The author wishes to thank Mr. L. H. Bainbridge-Bell, who developed the apparatus; Miss E. Rogers, for her assistance in the analysis of the records; and also Dr. R. L. Smith-Rose and Sir George Simpson, for suggestions in connection with the presentation of the paper.

REFERENCES

(1) R. A. Watson Watt: *Journal I.E.E.*, 1926, vol. 64, p. 596.
(2) R. A. Watson Watt, J. F. Herd, and L. H. Bainbridge-Bell: "The Cathode-Ray Oscillograph in Radio Research," pp. 145–172 (H.M. Stationery Office, 1933).
(3) R. A. Watson Watt and J. F. Herd: *Journal I.E.E.*, 1926, vol. 64, p. 611.
(4) C. E. P. Brooks: *Memoirs of the Royal Meteorological Society*, 1925, vol. 3, p. 145.
(5) R. A. Watson Watt: *Journal of the Royal Aeronautical Society*, 1925, vol. 29, p. 62.
(6) R. A. Watson Watt: *Quarterly Journal of the Royal Meteorological Society*, 1929, vol. 55, p. 274.
(7) R. Bureau: "Les Foyers d'Atmosphériques" (French Meteorological Office).
(8) G. H. Munro, W. J. Wark, and A. J. Higgs: Radio Research Board, Australia, 1936, Bulletin No. 100.
(9) Popoff: *Journal of the Russian Physico-chemical Society*, 1895, vol. 28, p. 899.
(10) M. J. Fényi: *Comptes Rendus*, 1902, vol. 134, p. 227.
(11) C. J. De Groot: *Proceedings of the Institute of Radio Engineers*, 1917, vol. 5, p. 75.
(12) F. Schindelhauer: *Jahrbuch der drahtlösen Telegraphie*, 1918, vol. 12, p. 153.
(13) R. A. Watson Watt: *Proceedings of the Royal Society*, A, 1922, vol. 102, p. 460.
(14) F. Schindelhauer: *Elektrische Nachrichten-technik*, 1932, vol. 9, p. 41.
(15) H. Hubert: *Comptes Rendus*, 1926, vol. 183, p. 368.
(16) L. W. Austin: *Proceedings of the Institute of Radio Engineers*, 1926, vol. 14, p. 373.
(17) M. Bäumler: *Jahrbuch der drahtlösen Telegraphie*, 1922, vol. 20, p. 456.
(18) C. T. R. Wilson: *Philosophical Transactions of the Royal Society*, A, 1920, vol. 221, p. 73.
(19) E. V. Appleton: *Proceedings of the Physical Society*, 1924, vol. 37, p. 48D.
(20) E. V. Appleton: Royal Society of Arts, Cobb Lectures, 1936.
(21) H. J. Round, T. L. Eckersley, K. Tremellen, and F. C. Lunnon: *Journal I.E.E.*, 1925, vol. 63, p. 933.

DISCUSSION ON
"THEORY AND PERFORMANCE OF THE ICONOSCOPE"*

Dr. F. C. Williams (*communicated*): The sections of this paper of greatest interest to me are those dealing with the signal/noise ratio. The exact calculation of noise in such a complicated device is a matter of considerable difficulty, and the authors appear to have made some simplification in the earlier section entitled " Calculation of the sensitivity of the iconoscope." Here they calculate the noise output with reference only to the thermal noise in the coupling resistor and shot noise in the amplifier; they omit the shot noise originating in the iconoscope.

They state on page 67 that "$\overline{V_N^2} = K_1 f R$.

Although the beam current does not in fact traverse the resistance R, it appears probable that its shot fluctuations will generate a noise across R relevant to such traversal. Further, the secondary and photo-currents which make up the actual return path of the beam current will probably contribute a further, more or less independent, shot fluctuation across R. Thus if I_b is the beam current, it appears very probable that there will be a fluctuation current traversing R whose mean-square value is

$$\overline{i^2} = 2 A_b I_b e d f$$

where A_b is some factor, probably not less than 2, which expresses the extent to which the fluctuations in R exceed those relevant to the passage of a temperature-limited current I_b through R. That A_b will be of the order of 2 seems to be implied in the section relating to electron multipliers; experimental evidence of its probability can be derived from the known properties of tetrode valves. With many such there exists an equilibrium potential for the anode (otherwise than zero) for which the anode current is zero. If the valve be set to operate at this point (by momentarily raising the anode potential above that of the screen) with a condenser and resistance in series connecting the anode and cathode, conditions are closely analogous to those obtaining in the iconoscope. Mr. E. B. Moullin has examined the fluctuations generated in R under such circumstances.† He found A_b to be of the order of 2 or more; precise statement is impossible since I_b was unknown and flicker effect was present.

It follows that the signal/noise ratio in the iconoscope is the same as that obtaining in a simple photocell and amplifier having a photo-current $A_b I_b$, the signal photo-current being given an arbitrary r.m.s. value I_0 for the time being. I have discussed this last circuit in a paper recently submitted to The Institution.‡ It is there shown that with the circuit of Fig. A the (noise/signal)^2 ratio is

$$S^2 = \frac{2 A_b I_b e f_0}{I_0^2}\left\{1 + \frac{2kT}{e R A_b I_b} + \frac{AI}{g^2 A_b I_b}\left(\frac{1 + \frac{1}{3}R^2 C^2 \omega_0^2}{R^2}\right)\right\}$$

where

$A_b I_b$ = mean photo-current.
I_0 = r.m.s. signal current from the photocell.
R = coupling resistance in series with the photocell.
C = interelectrode and stray capacitance from grid to earth.
I = mean anode current of the valve.
A = measure of the depression of shot effect by space charge.
g = mutual conductance of the valve.
$f_0 = \omega_0/2\pi$ = band width to which the amplifier responds;

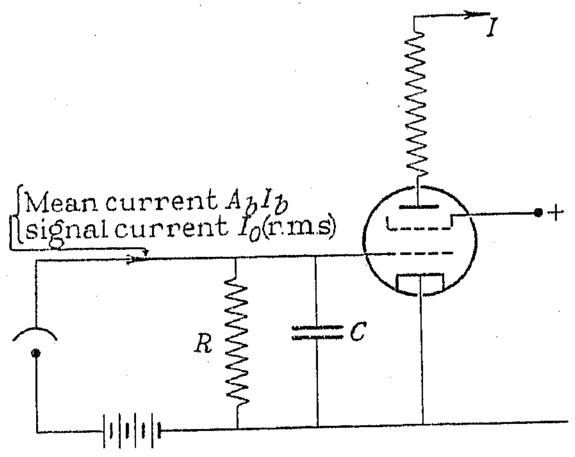

Fig. A

it is assumed that by means of correcting circuits later in the amplifier the response is held constant at all frequencies from zero to f_0 whatever the value of R, and that the response is zero above f_0.

It is evident that, as R is increased, S decreases steadily to the limit

$$S^2 = \frac{2 A_b I_b e f_0}{I_0^2}\left\{1 + \frac{A I C^2 \omega_0^2}{3 g^2 A_b I_b}\right\}$$

thermal noise in R, represented by the second term in the first expression, being then negligible. The term outside the bracket is the (noise/signal)^2 in the photo-current, or, in the case of the iconoscope, the basic limit at the mosaic. The second term inside the bracket represents the fractional extent to which valve noise supplements the basic noise. Valve noise will be negligible if

$$\frac{A I C^2 \omega_0^2}{3 g^2 A_b I_b} < 1$$

i.e.

$$I_b > \frac{A I C^2 \omega_0^2}{3 g^2 A_b}$$

---

* Paper by V. K. Zworykin, Ph.D., G. A. Morton, Ph.D., and L. E. Flory (see page 63). Reprinted from *Journal I.E.E.*, 1938, vol. 82, p. 561.

† *Proceedings of the Royal Society*, 1934, A, vol. 147, p. 100. Actually the condenser was short-circuited in his experiments. Since I is zero this is immaterial; the fluctuation in R will be unaffected by its introduction, provided it is great compared with the anode-cathode capacitance.

‡ " Co-existent Thermal and Thermionic Fluctuations in Complex Networks."

With typical values this reduces to

$$I_b > 0 \cdot 15\ \mu\text{A}.$$

Hence with a high value of coupling resistance ($R = 10^5$ to 10^6 ohms is sufficient) and subsequent frequency compensation, we shall have, if $I_b = 0 \cdot 5\ \mu\text{A}$,

$$S^2 \simeq \frac{2A_b I_b e f_0}{I_0^2}$$

With the authors' notation, and with $I_0 = FpkA$, this reduces to

$$\frac{1}{N} = S = \frac{\sqrt{(2A_b I_b e f_0)}}{FpkA}$$

or

$$F = \frac{N\sqrt{(2A_b I_b e f_0)}}{pkA}$$

Hence with $N = 10$, $A_b = 2$, and $I_b = 0 \cdot 5\ \mu\text{A}$, and their values for the remaining factors,

$$F = 0 \cdot 2 \times 10^{-3}\ \text{lumen/cm}^2$$

as against $0 \cdot 7 \times 10^{-3}$ lumen/cm² according to the authors.

Thus the use of high R and subsequent correction should improve the sensitivity by a factor of $3\frac{1}{2}$.

The improvement obtained by this method does not appear capable of attaining such values as are possible with secondary multiplication, for valves are not available for which $\dfrac{AIC^2\omega_0^2}{g^2Ab}$ is sufficiently low to permit of the advantageous use of much lower values of I_b. But it does appear that the adoption of the system would permit of the more advantageous use of the present commercially-produced iconoscopes.

It is perhaps worth noting that it can be shown from the first equation that the authors' assessment of valve noise is probably rather high; if the same typical valve properties are assumed for $R = 10^4\omega$ as were assumed for high values of R, the value of F is about $0 \cdot 5$ lumen/cm².

MEETINGS OF THE WIRELESS SECTION

139TH MEETING OF THE WIRELESS SECTION, 1ST DECEMBER, 1937

Mr. T. Wadsworth, M.Sc., Chairman of the Section, took the chair at 6 p.m.

The minutes of the meeting held on the 3rd November, 1937, were taken as read and were confirmed and signed.

A paper by Messrs. A. E. Barrett, M.A., and C. J. F. Tweed, Associate Members, entitled "Some Aspects of Magnetic Recording and its Application to Broadcasting" (see page 73), was read and discussed.

A vote of thanks to the authors, moved by the Chairman, was carried with acclamation.

140TH MEETING OF THE WIRELESS SECTION, 5TH JANUARY, 1938

Mr. T. Wadsworth, M.Sc., Chairman of the Section, took the chair at 6 p.m.

The minutes of the meeting held on the 1st December, 1937, were taken as read and were confirmed and signed.

A paper by Mr. D. A. Bell, B.A., B.Sc., Graduate, entitled "A Theory of Fluctuation Noise" (see page 97), was read and discussed.

A vote of thanks to the author, moved by the Chairman, was carried with acclamation.

12TH INFORMAL MEETING OF THE WIRELESS SECTION, 25TH JANUARY, 1938

Mr. T. Wadsworth, M.Sc., Chairman of the Section, took the chair at 6.30 p.m.

The minutes of the Informal Meeting held on the 23rd November, 1937, were taken as read and were confirmed and signed.

A discussion, opened by Mr. L. H. Bedford, M.A., took place on "Cathode-Ray Tubes for Wireless Purposes."

At the conclusion of the discussion a vote of thanks was accorded to Mr. Bedford for his introductory remarks.

141ST MEETING OF THE WIRELESS SECTION, 2ND FEBRUARY, 1938

Mr. W. J. Picken, Vice-Chairman of the Section, took the chair at 6 p.m.

The minutes of the meeting held on the 5th January, 1938, were taken as read and were confirmed and signed.

A paper by Prof. C. L. Fortescue, O.B.E., M.A., Member, and Mr. G. Mole, Ph.D., B.Sc., entitled "A Resonance Bridge for Use at Frequencies up to 10 Megacycles per Second" (see page 112), was read and discussed.

A vote of thanks to the authors, moved by Mr. Picken, was carried with acclamation.